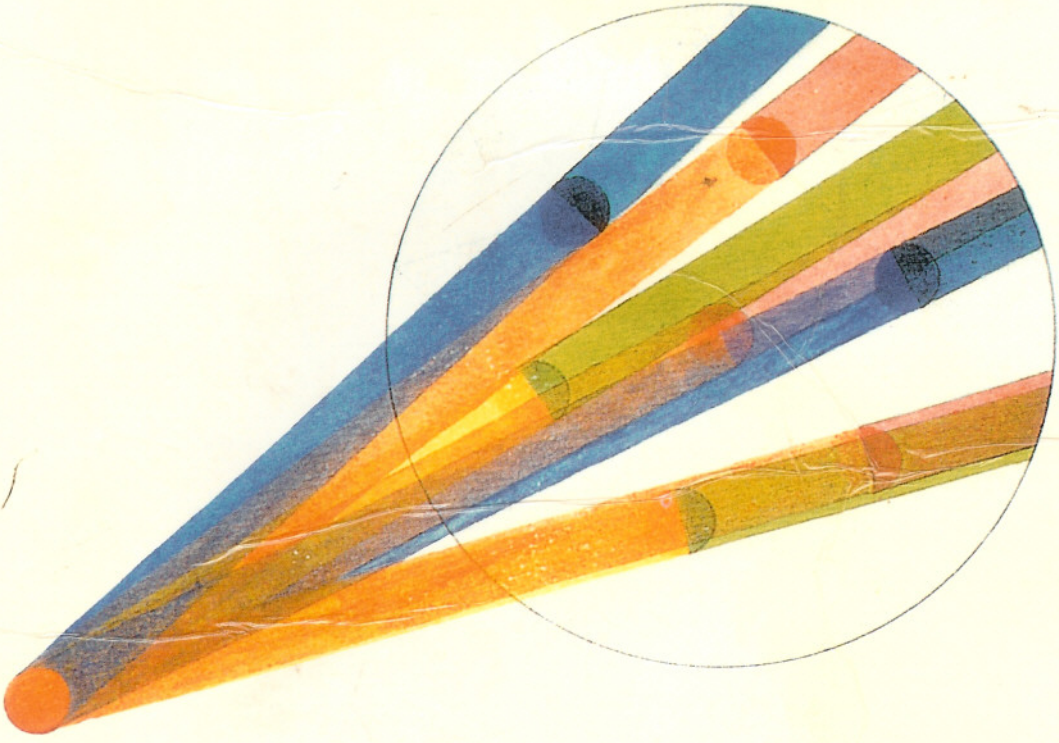
Marco Luise
Giorgio M. Vitetta



Teoria dei segnali

McGraw-Hill

VISIT…

Teoria dei segnali

Questo libro si propone al lettore come percorso di apprendimento guidato, piuttosto che come riferimento esaustivo sull'argomento; in questo senso, il volume aspira a colmare un vuoto significativo nell'attuale bibliografia in lingua italiana.
Il testo presenta con un approccio sostanzialmente unitario, e a pari dignità, i segnali determinati analogici (a tempo continuo) e digitali (a tempo discreto); inoltre, vista l'importanza crescente dei segnali digitali, vengono introdotti concetti tradizionalmente ritenuti di pertinenza dell'elaborazione numerica dei segnali (interpolazione e decimazione, filtri numerici, FFT...).
L'analisi di Fourier per i segnali periodici e aperiodici, a tempo continuo e discreto, viene applicata anche allo studio dei sistemi lineari stazionari monodimensionali.
I capitoli conclusivi sono dedicati a uno studio elementare dei segnali aleatori; a tal fine vengono richiamate le necessarie nozioni di probabilità e statistica.
L'esposizione, senza trascurare il rigore matematico, privilegia piuttosto gli aspetti intuitivi, anche con l'ausilio di numerosi esempi svolti ed esercizi.

Marco Luise è professore associato di Comunicazioni ottiche presso la Facoltà di Ingegneria dell'Università di Pisa.
Giorgio Matteo Vitetta è professore associato di Sistemi di telecomunicazione presso la Facoltà di Ingegneria dell'Università di Modena.

ISBN 88-386-0809-1

lire 54.000 (i.i.)

Indice

CENTRO BIBLIOTECARIO INGEGNERIA PISA

Prefazione xi

1 Introduzione allo studio dei segnali

1.1 Che cos'è un segnale? 1
1.2 Tipi di segnali 3
1.3 Proprietà elementari dei segnali determinati 9
Sommario 12
Esercizi proposti 13

2 Segnali periodici a tempo continuo

2.1 Dall'analisi fasoriale all'analisi di Fourier 15
2.2 Analisi armonica dei segnali periodici 18
2.2.1 Sviluppo in serie di Fourier in forma reale polare 18
2.2.2 Sviluppo in serie di Fourier in forma complessa 20
2.2.3 Sviluppo in serie di Fourier in forma reale rettangolare 22
2.3 Il criterio di Dirichlet 22
2.4 Spettri di ampiezza e di fase 24
2.5 Proprietà dello spettro di un segnale reale periodico 28
2.6 Segnali pari, dispari e alternativi 32
2.7 Sintesi del segnale con un numero limitato di armoniche 38
Sommario 47
Esercizi proposti 47

3 Segnali aperiodici a tempo continuo

3.1 Dalla serie all'integrale di Fourier 51
3.2 Proprietà della trasformata di Fourier 60
3.2.1 Criteri di esistenza 60
3.2.2 Simmetrie degli spettri 62
3.2.3 Segnali pari e dispari 63
3.3 Teoremi sulla trasformata di Fourier 63
3.3.1 Teorema di linearità 63
3.3.2 Teorema di dualità 64

3.3.3 Teorema del ritardo 65
3.3.4 Teorema del cambiamento di scala 68
3.3.5 Teorema della modulazione 73
3.3.6 Teorema di derivazione e integrazione 77
3.3.7 Teorema del prodotto 85
3.3.8 Teorema della convoluzione 86
3.4 Trasformate di Fourier generalizzate 93
3.4.1 La funzione generalizzata impulsiva o δ di Dirac 93
3.4.2 Proprietà della funzione generalizzata $\delta(t)$ 100
3.4.3 Trasformata di Fourier della funzione δ 103
3.4.4 Una trasformata notevole: la funzione $1/t$ 104
3.4.5 Trasformata della funzione gradino; teorema d'integrazione completo 106
3.4.6 Trasformata delle funzioni seno, coseno e dei segnali periodici 108
3.5 Periodicizzazione e formule di somma di Poisson 114
3.6 Relazione fra le trasformate di Laplace e di Fourier 117
Sommario 122
Esercizi proposti 123

Appendice: Cenni alla teoria delle distribuzioni

A.1 Definizione di distribuzione e funzione generalizzata 127
A.2 La funzione generalizzata di Dirac 129
A.3 Derivata di una distribuzione e di una funzione generalizzata 130
A.4 Limite di una distribuzione e di una funzione generalizzata 132

4 Sistemi monodimensionali a tempo continuo

4.1 Caratterizzazione dei sistemi a tempo continuo 133
4.1.1 Dal concetto di segnale al concetto di sistema 133
4.1.2 Proprietà dei sistemi monodimensionali 135
4.2 Caratterizzazione e analisi dei sistemi lineari stazionari 139
4.2.1 La risposta impulsiva 139
4.2.2 La risposta in frequenza 143
4.2.3 Il decibel 149
4.2.4 Sistemi in cascata e in parallelo 160
4.3 Filtri 162
4.3.1 Generalità sui filtri e filtri ideali 162
4.3.2 Criterio di Paley-Wiener e filtri reali 169
4.3.3 Banda e durata di un segnale e banda di un sistema 171
4.3.4 Distorsioni introdotte dai filtri 180
4.4 Densità spettrale di energia e potenza 186
4.4.1 Teorema di Parseval e densità spettrale di energia 186
4.4.2 Densità spettrale di potenza 189
4.4.3 Funzione di autocorrelazione e teorema di Wiener-Khintchine 191
4.4.4 Densità spettrale di potenza dei segnali periodici 198
4.5 Sistemi non lineari 200
4.5.1 Caratterizzazione dei sistemi non lineari 200
4.5.2 Nonlinearità essenziali e parassite 203
4.5.3 Misura delle distorsioni non lineari 205
Sommario 211
Esercizi proposti 213

5 Segnali a tempo discreto

5.1 Dal tempo continuo al tempo discreto 219
5.1.1 Campionamento dei segnali a tempo continuo 219
5.1.2 Alcuni segnali notevoli 222
5.2 Rappresentazione dei segnali aperiodici a tempo discreto nel dominio della frequenza 225
5.2.1 Trasformata di Fourier di una sequenza 225
5.3 Teoremi sulla trasformata di Fourier di una sequenza 233
5.3.1 Teorema di linearità 233
5.3.2 Teorema del ritardo 233
5.3.3 Teorema della modulazione 234
5.3.4 Teorema della somma di convoluzione 234
5.3.5 Teorema del prodotto 235
5.3.6 Teorema dell'incremento 236
5.3.7 Teorema della sequenza somma 236
5.4 La condizione di Nyquist e il teorema del campionamento 237
5.4.1 La condizione di Nyquist 237
5.4.2 Interpolazione a mantenimento 244
5.4.3 Interpolazione cardinale - Il teorema del campionamento 251
5.5 Analisi di Fourier delle sequenze periodiche 257
5.5.1 Trasformata discreta di Fourier 257
5.5.2 Teorema del prodotto 264
5.5.3 Teorema della convoluzione 264
5.5.4 Periodicizzazione di una sequenza aperiodica 267
5.6 Cenno agli algoritmi veloci di trasformata discreta (FFT) 269
5.6.1 Complessità di calcolo della trasformata discreta 269
5.6.2 Applicazioni dell'algoritmo di FFT: analisi spettrale 274
5.6.3 Applicazioni dell'algoritmo FFT: convoluzione veloce 277
5.7 Riassunto delle caratteristiche delle trasformate di Fourier 282
Sommario 283
Esercizi proposti 284

6 Sistemi monodimensionali a tempo discreto

6.1 Caratterizzazione dei sistemi a tempo discreto 289
6.1.1 Proprietà dei sistemi monodimensionali a tempo discreto 290
6.1.2 Sistemi lineari e stazionari a tempo discreto 291
6.1.3 Risposta in frequenza di un SLS 294
6.1.4 Filtri a tempo discreto 296
6.2 Cambiamento della frequenza di campionamento 303
6.2.1 Sovracampionamento con interpolazione numerica 303
6.2.2 Decimazione o sottocampionamento 312
6.3 Cenni alla trasformata $\mathcal{Z}$ di una sequenza 317
6.3.1 Definizione di trasformata $\mathcal{Z}$ e zone di convergenza 317
6.3.2 Relazione con la trasformata di Fourier 324
6.3.3 Inversione della trasformata $\mathcal{Z}$ 326
6.3.4 Proprietà della trasformata $\mathcal{Z}$ 326
6.4 Sistemi a tempo discreto regolati da equazioni alle differenze 328
6.4.1 Un caso di studio 328
6.4.2 Implementazione con componenti elementari e generalizzazione 331
6.4.3 Calcolo della risposta impulsiva 334

6.4.4 La funzione di trasferimento 337
6.4.5 Sistemi a risposta impulsiva finita e infinita 338
6.5 Cenni al progetto di filtri numerici IIR 342
6.5.1 La tecnica dell'invarianza impulsiva 343
6.5.2 La tecnica della trasformazione bilineare 350
Sommario 359
Esercizi proposti 361

7 Richiami di teoria della probabilità

Premessa 369
7.1 Esperimenti deterministici e aleatori 369
7.2 Elementi di teoria della probabilità
7.2.1 Esperimento aleatorio, spazio di probabilità e proprietà elementari 371
7.2.2 Esperimento aleatorio composto 378
7.3 Variabili aleatorie 381
7.3.1 Definizione di variabile aleatoria 381
7.3.2 Densità di probabilità di una variabile aleatoria 384
7.3.3 Trasformazione di una variabile aleatoria 390
7.3.4 Indici caratteristici di una distribuzione 393
7.3.5 La variabile aleatoria Gaussiana 398
7.3.6 Variabili aleatorie condizionate 401
7.4 Sistemi di variabili aleatorie 404
7.4.1 Sistemi di due variabili aleatorie 404
7.4.2 Funzioni distribuzione e densità di probabilità condizionate 407
7.4.3 Trasformazione di una coppia di variabili aleatorie 408
7.4.4 Correlazione e covarianza 412
7.4.5 Sistemi di *n* variabili aleatorie e vettori aleatori 414
7.4.6 Trasformazione di un vettore aleatorio 416
7.4.7 Variabili aleatorie congiuntamente Gaussiane (vettori Gaussiani) 420
7.4.8 Il teorema-limite centrale 423
Sommario 426
Esercizi proposti 426

8 Segnali aleatori

8.1 Dai segnali determinati ai segnali aleatori 431
8.1.1 Definizione di processo aleatorio 433
8.1.2 Processi parametrici 435
8.1.3 Caratterizzazione statistica di un processo aleatorio 437
8.2 Indici statistici del 1° e 2° ordine di un processo aleatorio 441
8.2.1 Funzioni valor medio, potenza, varianza 441
8.2.2 Funzioni di autocorrelazione e autocovarianza 444
8.3 Processi aleatori stazionari 446
8.3.1 Stazionarietà in senso stretto 446
8.3.2 Stazionarietà in senso lato 449
8.3.3 Proprietà della funzione di autocorrelazione di un processo stazionario in senso lato 455
8.4 Filtraggio di un segnale aleatorio 461
8.4.1 Relazione ingresso-uscita tra le statistiche semplificate 461
8.4.2 Filtraggio di un processo aleatorio stazionario in senso lato 465
8.5 Densità spettrale di potenza di un processo stazionario 467

8.5.1 Definizione e teorema di Wiener-Khintchine 467
8.5.2 Filtraggio di un processo aleatorio e densità spettrale di potenza 470
8.5.3 Processo di rumore bianco 472
8.6 Processi aleatori Gaussiani 482
8.6.1 Definizione e prime proprietà 482
8.6.2 Filtraggio dei processi Gaussiani 484
8.7 Processi ergodici 493
8.7.1 Ergodicità del valore medio 493
8.7.2 Ergodicità della funzione di autocorrelazione - Ergodicità in senso stretto 498
Sommario 501
Esercizi proposti 503
Indice analitico 511

1

Introduzione allo studio dei segnali

1.1 Che cos'è un segnale?

La definizione di *segnale* non è immediata. Esempi familiari tratti dalla vita quotidiana sono il segnale acustico prodotto da uno strumento musicale (che dal punto di vista fisico può essere caratterizzato come una variazione della pressione dell'aria provocata dallo strumento, e rilevata dal nostro orecchio); il segnale misurato da un elettrocardiografo (una debole tensione elettrica) e registrato sulla tipica "strisciata"; il segnale radio (un campo elettromagnetico variabile) captato dall'antenna di un ricevitore; il segnale luminoso emesso da una lampadina di un semaforo, o da un apparecchio televisivo, e così via. Tutti gli esempi precedenti hanno in comune una caratteristica, e cioè il fatto che il segnale esiste in quanto si fa *portatore di una informazione* che giustifica l'esistenza e l'importanza del segnale stesso. Questa informazione può essere di varia natura: di carattere estetico, nel caso del brano musicale, medico, nel caso dell' elettrocardiogramma, e così via.

Tentando dunque di sintetizzare quanto sopra, possiamo dire che un segnale è *una qualunque grandezza fisica variabile cui è associata una informazione*. In molti casi, l'andamento del segnale può essere perfettamente noto, ad esempio attraverso una registrazione su carta (come per il caso dell'elettrocardiogramma), su nastro magnetico o come un *file* in un calcolatore elettronico. Dunque il modo più conveniente per caratterizzare, studiare ed elaborare un segnale passa attraverso la schematizzazione dello stesso come una *funzione matematica di una o più variabili*. L'elettrocardiogramma rappresentato in

Figura 1.1 può essere considerato come *il grafico di una funzione di variabile reale a valori reali* $x(t): \Re \rightarrow \Re$ ove la variabile indipendente t ha il significato di un *tempo*, e il valore del segnale $x(t)$ rappresenta l'andamento della tensione raccolta dall'apparato biomedicale. La notazione usata riflette questo caso tipico, in cui cioè l'evoluzione del segnale avviene in ambito *temporale*.

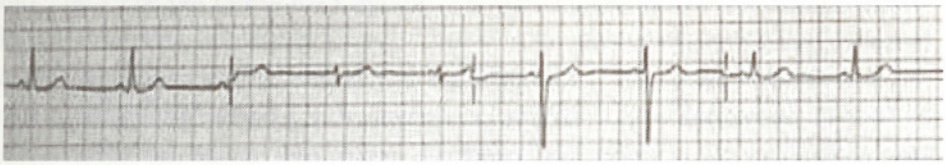


Figura 1.1 Esempio di elettrocardiogramma

L'esempio precedente di segnale monodimensionale *nel tempo* non esaurisce ovviamente i diversi tipi di segnale con cui si può avere a che fare. La semplice immagine in bianco e nero in Figura 1.2 deve essere considerata anch'essa un *segnale*, poiché, concordemente con la definizione appena data, ha in sé una certa *informazione*. La schematizzazione appropriata è stavolta quella di una *funzione bidimensionale* $z(x_1,x_2): \Re^2 \rightarrow \Re$ delle due coordinate spaziali (x_1,x_2) (si veda la Figura 1.2), ove il valore z rappresenta l'*intensità luminosa* del generico punto sull'immagine (detto nel gergo dell'elaborazione delle immagini *pixel*) di coordinate (x_1,x_2). La *variazione* del segnale avviene in un ambito di tipo *spaziale*, nel senso che si hanno diversi valori del segnale (l'intensità luminosa) per diversi valori delle coordinate spaziali del generico pixel.

Se consideriamo inoltre una immagine *a colori,* essa può essere rappresentata dalla sovrapposizione di tre distinte immagini nei cosiddetti colori fondamentali Rosso (R, Red), Verde (G, Green) e Blu (B, Blue). Ciascuna di queste immagini è caratterizzata da un diverso andamento della rispettiva intensità luminosa sui vari pixel. In questo caso abbiamo a che fare con un segnale bidimensionale *vettoriale* le cui componenti sono rispettivamente le tre intensità dei canali RGB: $\mathbf{z}(x_1,x_2) = \left[z_R(x_1,x_2) \quad z_G(x_1,x_2) \quad z_B(x_1,x_2)\right]$.

Come ulteriore esempio, la Figura 1.3 rappresenta una moltitudine di segnali sismici rilevati a un dato istante da vari sensori in diversi siti, segnali che devono essere considerati congiuntamente come un segnale vettoriale $\mathbf{z}(t)$ per ottenere la massima informazione sullo stato geofisico del territorio.

Per una introduzione allo studio dei segnali, è sufficiente trattare soltanto segnali monodimensionali di una variabile che, salvo diversa indicazione, dovrà intendersi di carattere temporale.

Figura 1.2 Esempio di segnale bidimensionale

1.2 Tipi di segnali

Una prima classificazione dei segnali può essere fatta proprio in base ai valori assunti dalla variabile indipendente, che negli esempi del Paragrafo 1.1 abbiamo per semplicità sempre supposto essere reale. Distinguiamo infatti tra:

- *segnali a tempo continuo*, per i quali il dominio della funzione ha la cardinalità dell'insieme dei numeri reali. La variabile indipendente può assumere *con continuità* tutti i valori compresi entro un certo intervallo, eventualmente illimitato. Il simbolo che useremo per la variabile temporale (continua) sarà t, e i segnali saranno indicati con $x(t)$, $y(t)$ ecc. L'elettrocardiogramma di Figura 1.1 e i vari segnali sismici in Figura 1.3 sono esempi tipici di segnali a tempo continuo;
- *segnali a tempo discreto*, per i quali il dominio della funzione ha la cardinalità dell'insieme (discreto) dei numeri interi. Tali segnali vengono chiamati in matematica *successioni* e indicati con il simbolo x_n, y_n ecc. Più specificamente, le successioni vengono chiamate nella teoria dei segnali *sequenze*, e saranno indicate con espressioni del tipo $x[n]$, $y[n]$, ove la variabile "temporale" n viene racchiusa tra parentesi quadre per evidenziarne l'intrinseca diversità dalla variabile (continua) t, che viceversa viene racchiusa tra parentesi tonde. Un segnale cinematografico, come è noto, è

ottenuto proiettando una *sequenza* di 24 fotogrammi (immagini) al secondo che dà all'occhio umano l'illusione di un segnale a tempo continuo (si veda a questo proposito *l'interpolatore con mantenimento* del Capitolo 5). Allora, come si suggerisce in Figura 1.4, il segnale cinematografico è una funzione tridimensionale $z(x_1, x_2, n]$ di due variabili spaziali continue che identificano i pixel dell'immagine, e di una ulteriore variabile temporale *discreta* n che identifica i vari fotogrammi in successione.

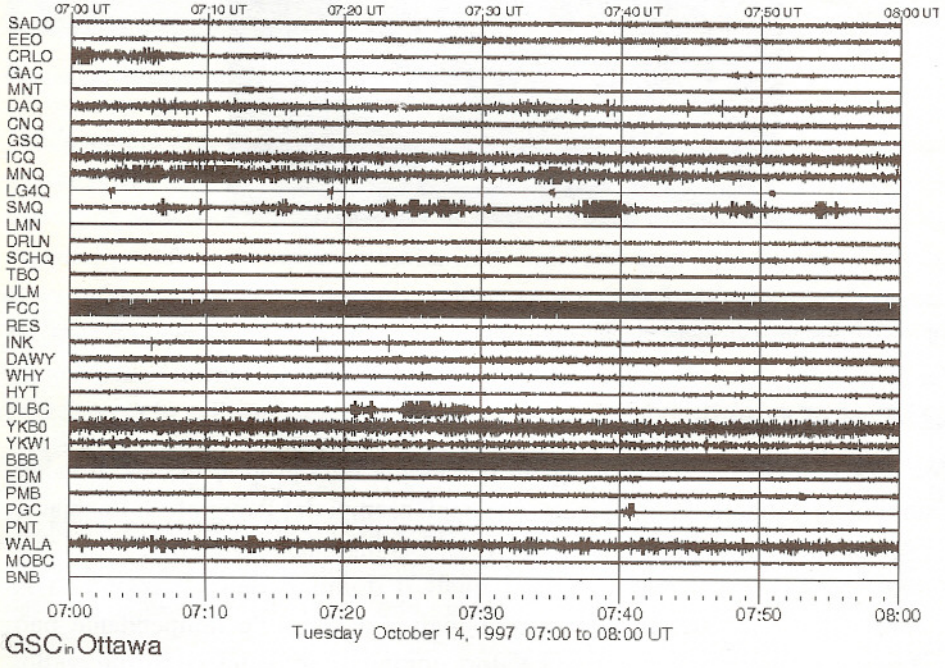


Figura 1.3 Registrazione di segnali sismici

Una classificazione analoga può essere condotta sulla base dei *valori* assunti dai segnali (cioè sulla base del codominio della funzione che li rappresenta). Avremo:

- *segnali ad ampiezza continua*, che possono assumere con continuità tutti i valori reali di un intervallo (eventualmente illimitato), come nel caso di un segnale acustico e in generale dei segnali osservati nei sistemi naturali;
- *segnali ad ampiezza discreta*, aventi come codominio un insieme numerabile, eventualmente illimitato. Il segnale luminoso prodotto da una lampadina di un semaforo può assumere solo due valori (acceso o spento), così come i segnali *binari* che regolano il funzionamento dei circuiti elettronici digitali.

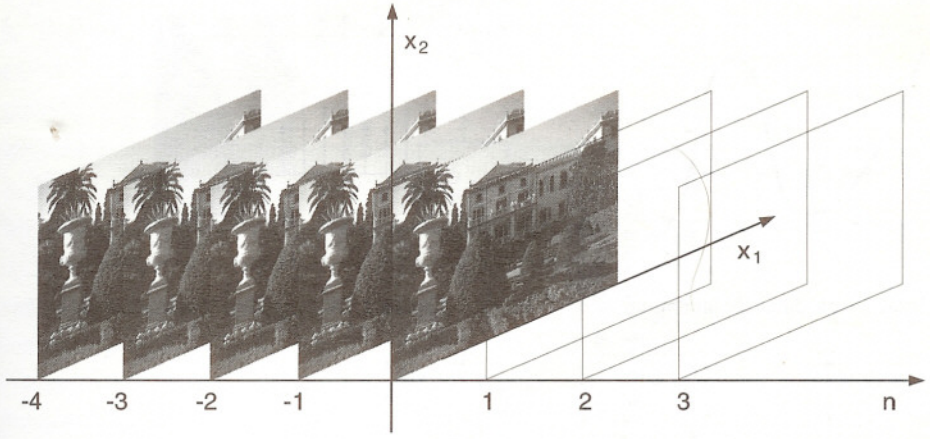


Figura 1.4 Segnale cinematografico a tempo discreto $z(x_1, x_2, n]$

I segnali a tempo continuo e ad ampiezza continua si dicono *analogici*, mentre quelli a tempo e ampiezza discreti si dicono *numerici* e, come già accennato, sono quelli tipicamente trattati dai calcolatori elettronici. Anche i segnali a tempo discreto e ampiezza continua (cioè le *sequenze* a valori reali) hanno una grande importanza perché costituiscono l'oggetto delle tecniche di elaborazione numerica dei segnali (DSP, Digital Signal Processing) che hanno avuto un enorme sviluppo negli ultimi trent'anni, come vedremo sommariamente nei Capitoli 4 e 5. Dei segnali ad ampiezze discrete e tempo continuo (talvolta detti *quantizzati*) non avremo più modo di discutere in seguito. Le Figure 1.5a-d mostrano esempi-tipo delle 4 classi di cui sopra; per meglio chiarire però la differente natura di questi diversi tipi di segnali, esaminiamo un esempio piuttosto familiare di applicazione delle tecniche di elaborazione dei segnali.

Esempio 1.1

La Figura 1.6 è uno schema semplificato di un sistema di registrazione di un segnale acustico su Compact-Disc (CD). Il segnale utile che deve essere registrato è la variazione di pressione acustica $p(t)$ prodotta dalla *sorgente* del segnale stesso, cioè il pianoforte. Tale segnale viene convertito in una debole tensione elettrica $v(t)$ dal microfono, che svolge la funzione di *trasduttore* (cioè di dispositivo che cambia la natura del segnale senza alterarne la forma). La tensione prodotta dal microfono, prima di poter essere ulteriormente elaborata, deve essere *amplificata* dal dispositivo amplificatore indicato con il simbolo triangolare.

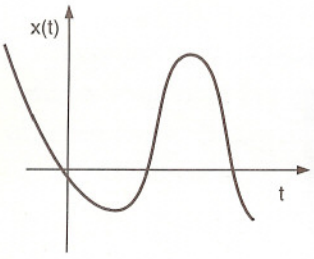


Figura 1.5a Segnale analogico

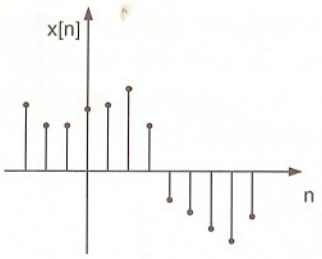


Figura 1.5b Segnale a tempo discreto

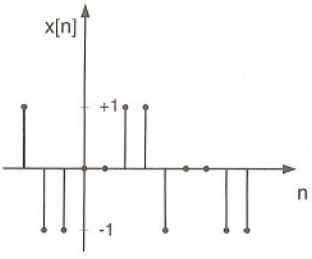


Figura 1.5c Segnale numerico

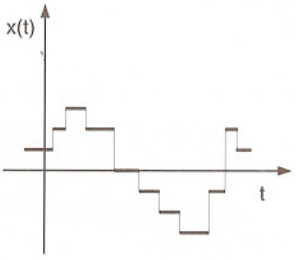


Figura 1.5d Segnale quantizzato

All'uscita di tale amplificatore ideale troviamo la tensione $x(t) = a \cdot v(t)$, ove $a > 1$ è il coefficiente di amplificazione. Sia il microfono sia l'amplificatore sono componenti *analogici* poiché trattano segnali analogici. Infatti l'andamento di $v(t)$, se il microfono non introduce *distorsioni*, è *analogo* a quello del segnale originario $p(t)$, e così anche il segnale $x(t)$, a meno di una costante moltiplicativa, replica fedelmente l'andamento di $v(t)$.

Poiché però si desidera registrare il segnale con componenti e circuiti *digitali* dobbiamo compiere ulteriori operazioni per *adattare* il segnale al mezzo. In particolare, il funzionamento di tutto il sistema è regolato da un elaboratore (microprocessore) che sostanzialmente può essere considerato alla stregua di un complesso circuito digitale *sincrono*. Come è noto, il funzionamento di un circuito sincrono avviene secondo una *sequenza* di singoli passi successivi regolati da un segnale di temporizzazione detto *clock*. Un segnale la cui variazione avviene a tempo continuo è quindi intrinsecamente inadatto a essere trattato da un microprocessore funzionante per passi *discreti nel tempo*. La Figura 1.6 mostra che la forma d'onda $x(t)$ viene allora *campionata*, ottenendo la *sequenza* $x[n]$ dei valori di $x(t)$ considerati ai multipli di un opportuno *periodo di campionamento* T: $x[n] = x(nT)$.

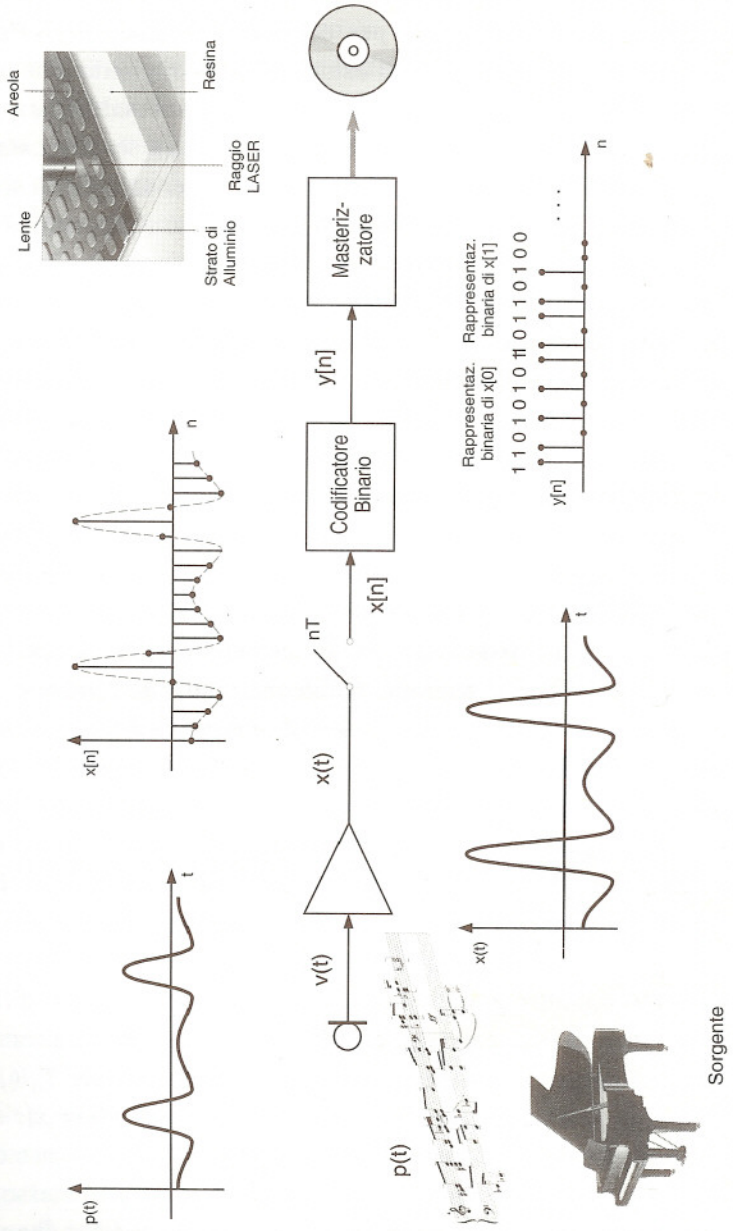


Figura 1.6 Schema di un sistema di registrazione su Compact Disc

Il segnale è stato adesso ridotto a tempo discreto, ma i vari *campioni* di $x[n]$ assumono ancora infiniti valori nell'insieme dei reali. Si deve quindi

ulteriormente procedere a una *codifica* di questi valori reali attraverso l'alfabeto *binario* tipico dei circuiti digitali, per ottenere il segnale *digitale binario* (numerico) $y[n]$ di Figura 1.6. Nell'esempio della registrazione su CD, il campionamento del segnale avviene alla cadenza standardizzata di 44100 campioni/secondo, e la codifica binaria dei valori reali del segnale campionato è in virgola fissa su 16 bit (in Figura 1.6 sono indicati per semplicità soli 8 bit di codifica). L'unione delle operazioni di campionamento e di codifica prende il nome di *conversione analogico/digitale*, e viene realizzata da appositi circuiti elettronici detti appunto *convertitori A/D*.

Il segnale digitale binario $y[n]$ viene quindi registrato sul CD dal cosiddetto *masterizzatore*. Si noti che il segnale temporale binario derivato in ultima analisi da $p(t)$ è registrato sul CD come un segnale *spaziale*: i valori delle cifre binarie (*bit*) di $y[n]$ che si susseguono nel tempo vengono registrati sotto forma di areole riflettenti o assorbenti la luce (a seconda del valore 0 o 1) lungo un "solco" a spirale che si svolge dal centro verso la periferia del CD stesso. Questo segnale sarà poi riletto dal laser dell'apparecchio lettore (come indicato in Figura 1.6), riconvertito in segnale analogico, amplificato e inviato a un altoparlante per ricostruire con la massima fedeltà il segnale-messaggio originario $p(t)$.

Nei prossimi capitoli risponderemo ad alcune domande fondamentali che il lettore dovrebbe essersi già posto a proposito di questo esempio: cosa succede se il microfono e/o l'amplificatore non sono perfettamente *fedeli*? Quando si effettua un campionamento, sotto quali condizioni non viene persa l'*informazione* del segnale analogico di partenza? Perché l'intervallo di campionamento T nel sistema CD è pari a 1/44100 di secondo? Viene forse falsata la natura del segnale nella codifica binaria dei valori reali? ❑

Distingueremo poi altre grandi classi di segnali secondo criteri differenti da quelli appena esposti. Nel caso di segnali a tempo continuo, diremo che un segnale è *periodico* quando esiste un certo intervallo temporale T_0 (N_0 per un segnale a tempo discreto) tale che si possa scrivere $x(t) = x(t + T_0)$ per qualunque generico valore del tempo t ($x[n] = x[n + N_0]$ $\forall n$ nel caso del tempo discreto). Ciò significa che il segnale si "ripete" uguale a se stesso dopo un periodo di tempo T_0 (N_0). Se non esiste alcun valore di T_0 (N_0) che verifica la relazione suddetta, il segnale è non periodico o *aperiodico*.

Nella considerazione delle varie proprietà ed esempi visti finora, è stato implicitamente assunto che il valore del segnale fosse univocamente determinabile non appena fossero fissati i valori delle variabili indipendenti (in

particolare, il tempo per segnali monodimensionali). Questo accade quando il segnale è noto attraverso un grafico, o una registrazione magnetica, o più semplicemente attraverso una ben definita espressione matematica, o ancora perché è il prodotto di sistemi e apparati di cui si ha stretto controllo (ad esempio, un generatore di forme d'onda di un laboratorio elettronico). In tali casi, diremo che il segnale è *determinato* o *deterministico*.

Viceversa, in moltissimi casi non è possibile conoscere con esattezza *a priori* il valore assunto da un segnale in un certo istante. Si pensi al segnale geofisico raccolto da sensori posti sul terreno per effettuare rilevazioni minerarie. Tale segnale non è noto *a priori* completamente, in particolare non se ne conosce l'evoluzione futura se non dopo l'osservazione, cioè *a posteriori*. Prima dell'osservazione, si ha solo una conoscenza generica di alcune proprietà di massima di tale segnale, derivante dall'esperienza pregressa in casi simili. Stessa osservazione può farsi a proposito delle tensioni di disturbo (*rumore*) presenti nei componenti elettronici attivi e passivi e prodotte da fenomeni incontrollabili, tipicamente di origine quantistica. Diremo quindi che questi segnali sono *aleatori*, intendendo che il valore assunto da essi è affetto da un certo grado di imponderabilità (alea) che ne impedisce una conoscenza esatta. Mentre per modellare e studiare i segnali determinati sono sufficienti i concetti dell'analisi matematica tradizionale, per i segnali aleatori è indispensabile ricorrere a tecniche basate sulla *teoria della probabilità e dei processi aleatori*, che saranno oggetto dei Capitoli 7 e 8.

1.3 Proprietà elementari dei segnali determinati

Limitiamo per il momento la nostra attenzione ai segnali determinati e definiamo alcune grandezze di fondamentale importanza per il prosieguo dello studio.

Supponiamo di disporre di un resistore di resistenza R attraversato da una corrente $i(t)$; l' espressione della potenza istantanea dissipata sul resistore per effetto Joule è, come è noto, $Ri^2(t)$. Osserviamo quindi la proporzionalità tra la potenza istantanea e il *quadrato* del segnale; il coefficiente di proporzionalità è legato al particolare esempio. Estendendo in maniera astratta tale definizione, diremo che al segnale $x(t)$ è associata una potenza istantanea *normalizzata* (aggettivo che verrà poi sistematicamente omesso) pari a $x^2(t)$. Inoltre, tornando all'esempio del resistore, l'energia totale dissipata per effetto del passaggio della corrente $i(t)$ è pari a $\int_{-\infty}^{+\infty} Ri^2(t)dt$. Conseguentemente, definiremo l'*energia* associata al segnale $x(t)$ come

$$E_x \triangleq \int_{-\infty}^{+\infty} |x(t)|^2 dt \tag{1.3.1}$$

purché l'integrale risulti convergente (cioè $E_x < \infty$)[1]. La definizione di energia, benché meno intuitiva, viene banalmente estesa anche ai segnali a tempo discreto come segue:

$$E_x \triangleq \sum_{n=-\infty}^{+\infty} |x[n]|^2 \quad < \infty \tag{1.3.2}$$

Per tutti i segnali *fisici* (cioè effettivamente osservati) l'integrale (o la sommatoria) che definisce l'energia risulta convergente, poiché ogni segnale proveniente da un sistema fisico è portatore di *energia finita*. Molto spesso però conviene considerare *modelli ideali* di segnale, ovvero segnali idealizzati non esistenti in natura, ma assai utili per approssimare casi reali. La tensione $v(t)$ generata dalla "batteria ideale" in Figura 1.7 (cioè da un generatore ideale di tensione) è costante per ogni valore di t da $-\infty$ a $+\infty$ (e ciò è fisicamente assurdo). Se però la batteria "reale" viene osservata durante il periodo in cui è carica ed eroga la tensione nominale, $v(t)$ è un'ottima approssimazione del segnale reale. Chiaramente, $v(t)$ possiede *energia illimitata* (nel senso che $\int_{-\infty}^{+\infty} |v(t)|^2 \, dt = \infty$), e quindi la definizione (1.3.1) è per questo segnale *mal posta*.

Consideriamo dunque un generico segnale $x(t)$ a valori limitati ma energia infinita, e costruiamo il segnale $x_T(t)$ con una operazione di *troncamento* come segue:

$$x_T(t) = \begin{cases} x(t) & |t| \le T/2 \\ 0 & \text{altrove} \end{cases} \tag{1.3.3}$$

La Figura 1.8 illustra graficamente questo procedimento. Evidentemente, $x_T(t)$ è ottenuto limitando l'osservazione all'intervallo finito $-T/2 \le t \le T/2$. Se indichiamo con E_{x_T} l'energia di $x_T(t)$, è chiaro che in generale $E_{x_T} < \infty$ poiché il segnale è diverso da zero e assume valori finiti solo su di un intervallo limitato. Altrettanto chiaro è che se ingrandiamo l'intervallo di osservazione per comprendere l'andamento di tutto il segnale $x(t)$ (imponiamo cioè $T \to \infty$) otteniamo $E_{x_T} \to \infty$. Introduciamo allora il concetto di *potenza* di un segnale: la potenza media del segnale $x_T(t)$ valutata sull'intervallo di osservazione $[-T/2, T/2]$ è per definizione pari all'energia di $x_T(t)$ rapportata alla durata

[1] Il simbolo $\triangleq$ significa "per definizione uguale a".

dell'intervallo stesso: $P_{x_T} \triangleq E_{x_T} / T$. Siamo ora in grado di estendere a $x(t)$ questa definizione di *potenza media* attraverso un'operazione al limite:

$$P_x \triangleq \lim_{T\to\infty} P_{x_T} = \lim_{T\to\infty} \frac{E_{x_T}}{T} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 \, dt \tag{1.3.4}$$

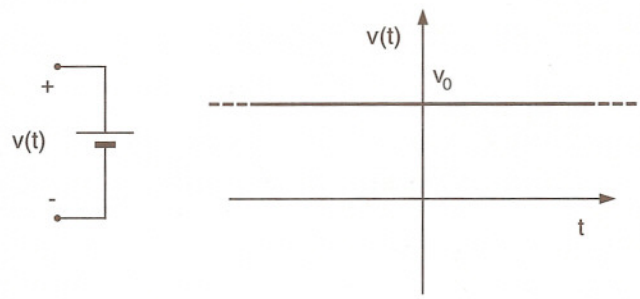


Figura 1.7 Segnale prodotto da una batteria ideale

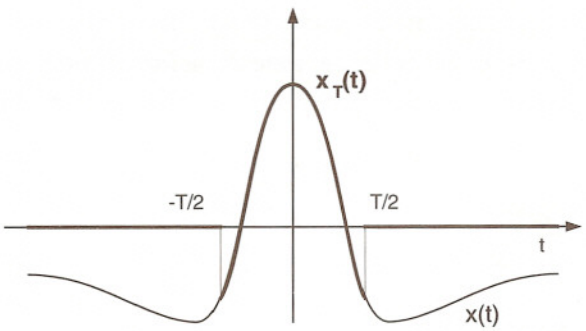


Figura 1.8 Troncamento del segnale $x(t)$

Esempio 1.2

Calcoliamo la potenza del segnale della batteria ideale di Figura 1.7. Dalla (1.3.4) si ha

$$P_v = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} |v(t)|^2 \, dt = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} v_0^2 \, dt = v_0^2 \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} 1 \cdot dt = v_0^2 \tag{E1.2.1}$$

come era ragionevole aspettarsi. ❑

Osserviamo che un segnale a energia finita (matematicamente, *a quadrato sommabile*) ha potenza media nulla; viceversa, un segnale che abbia un valore finito diverso da zero della potenza media ha necessariamente energia infinita.

Analogamente al caso della (1.3.2), per i segnali a tempo discreto abbiamo inoltre

$$P_x \triangleq \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} |x[n]|^2 \tag{1.3.5}$$

ove la notazione è autoesplicativa.

Talvolta torna utile usare il *valore efficace* di un segnale a potenza finita, definito sia per i segnali a tempo continuo, sia per quelli a tempo discreto come

$$x_{eff} \triangleq \sqrt{P_x} \tag{1.3.6}$$

Ricordiamo che il valore efficace di un dato segnale (chiamato nei paesi anglosassoni RMS, Root Mean Square) si può interpretare come quel valore che dovrebbe assumere un segnale costante per avere lo stesso contenuto in potenza del segnale dato.

Terminiamo il capitolo con la definizione di *valor medio temporale* di un segnale, che richiede un procedimento al limite simile a quello appena visto relativamente alla potenza:

$$x_m \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\,dt \tag{1.3.7a}$$

$$x_m \triangleq \lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} x[n] \tag{1.3.7b}$$

Nell'ingegneria elettrica, il valor medio rappresenta la "componente continua" (cioè costante) attorno alla quale si svolge l'evoluzione temporale del segnale.

1.4 Sommario

Dopo aver definito un segnale come una grandezza fisica variabile portatrice di una informazione, abbiamo individuato alcune grandi classi di segnali, a *tempo continuo* e a *tempo discreto*, e anche ad *ampiezza continua* e *ampiezza discreta*. I segnali a tempo e ampiezza continui si dicono *analogici*, quelli a tempo e ampiezza discreti si chiamano *numerici*. I segnali a tempo discreto e ampiezza continua sono le *sequenze* che hanno grande importanza nelle tecniche di elaborazione numerica del segnale (DSP). I segnali sono inoltre *determinati* se il

loro andamento è noto in anticipo (*a priori*) per tutti i valori del tempo, altrimenti sono *aleatori* (noti *a posteriori*). Sono inoltre *periodici* se si ripetono con regolarità dopo un intervallo temporale finito, altrimenti sono *aperiodici*. Le proprietà elementari dei segnali, e cioè *energia*, *potenza*, *valor medio*, rappresentano il punto di partenza per uno studio e un'analisi più approfondita che inizierà nel prossimo capitolo con i *segnali periodici a tempo continuo*.

Esercizi proposti

1.1 Un apparato di rilevazione automatica a un casello autostradale conta i veicoli in ingresso al casello e annota l'orario d'ingresso. Considerando questa serie di osservazioni come un *segnale*, dire di che *tipo* di segnale si tratta, nel senso della Figura 1.5.

1.2 Spiegare perché un elettrocardiogramma è un segnale periodico. Osservando poi che la frequenza cardiaca di un adulto maschio a riposo è all'incirca 60 battiti al minuto, trovare il periodo di ripetizione T_0 del segnale.

1.3 Spiegare perché un segnale è *aleatorio* solo finché non lo si è completamente osservato, dopo di che rientra anch'esso nella categoria dei segnali *determinati*. L'elettrocardiogramma di un paziente è un segnale aleatorio o determinato *prima* che venga misurato? E dopo?

1.4 Rappresentare il grafico del segnale

$$x(t) = \begin{cases} 0 & t < 0 \\ \exp(-t/T) & t \geq 0 \end{cases}$$

e calcolarne il valor medio x_m, l'energia E_x e la potenza P_x.

1.5 Calcolare il valor medio e la potenza dei seguenti segnali:

i) $x(t) = a \cdot \cos(2\pi t / T_0)$

ii) $x(t) = \sin(2\pi t / T_0)$

iii) $x(t) = \mathrm{sgn}(t) \triangleq \begin{cases} 1 & t > 0 \\ 0 & t = 0 \\ -1 & t < 0 \end{cases}$

iv) $x(t) = \mathrm{sgn}\left(a \cdot \cos(2\pi t / T_0)\right)$

v) $x(t) = \mathrm{u}(t) \triangleq \begin{cases} 1 & t > 0 \\ 1/2 & t = 0 \\ 0 & t < 0 \end{cases}$

1.6 Ripetere l'Esercizio 1.4 per il segnale $x[n] = \exp(-n/N),\ n \geq 0$.

1.7 Considerare il segnale

$$x(t) = \begin{cases} 1 & |t| < T/2 \\ \\ 0 & \textit{altrimenti} \end{cases}$$

e dire perché il problema di calcolarne il valore efficace x_{eff} è *mal posto.*

2

Segnali periodici a tempo continuo

2.1 Dall'analisi fasoriale all'analisi di Fourier

Introduciamo in questo paragrafo l'*analisi di Fourier* dei segnali periodici con un esempio tratto dalla teoria dei circuiti lineari. Come è noto, i metodi del calcolo fasoriale permettono di risolvere problemi anche di una certa rilevanza e di un certo grado di complessità relativi al *regime sinusoidale* dei circuiti *lineari a componenti concentrati e di valore costante nel tempo*. Un esempio semplicissimo di questa classe di circuiti è dato in Figura 2.1, che mostra il circuito passivo detto "squadra R-C". Rivediamo brevemente come viene applicato il calcolo fasoriale a questo problema. Supponiamo di fornire al circuito la tensione sinusoidale

$$x(t) = a \cdot \cos(2\pi f_0 t + \theta) \tag{2.1.1}$$

Come è noto, a è l'*ampiezza* della oscillazione, $f_0 = 1/T_0$ ne è la *frequenza* e θ ne è la *fase iniziale*. Se il *periodo* dell'oscillazione T_0 è misurato in secondi, allora la frequenza f_0 è misurata in cicli/secondo, ovvero *hertz* (Hz). La Figura 2.2a mostra un esempio di segnale sinusoidale con i particolari valori di a, f_0 e θ indicati nella didascalia.

Il *fasore* $\mathring{X}$ associato alla sinusoide $x(t)$ è per definizione quel numero complesso tale che

$$x(t) = \Re\left[\mathring{X}\exp(j2\pi f_0 t)\right] \tag{2.1.2}$$

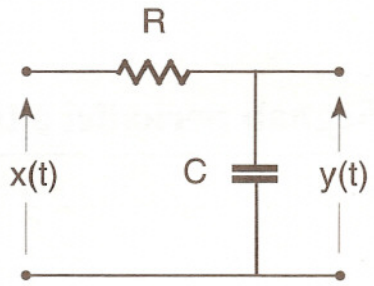


Figura 2.1 Squadra *R-C*

ove l'operatore $\Re[\cdot]$ indica "parte reale" e $j \triangleq \sqrt{-1}$ è l'unità immaginaria. Nel nostro caso, $\overset{\circ}{X} = a \exp(j\theta)$. A questo punto si desidera evidentemente ricavare l'andamento della tensione $y(t)$ ai capi del condensatore di capacità C in Figura 2.1. Le proprietà dei circuiti lineari a componenti costanti assicurano che anche la tensione $y(t)$ sarà sinusoidale alla stessa frequenza di $x(t)$. Del segnale cercato restano soltanto da determinare l'ampiezza e la fase iniziale.

Per far questo, applichiamo la legge del "partitore di tensione" al circuito dato. Tenendo conto che l'impedenza del condensatore alla frequenza data è $z_C = 1/(j2\pi f_0 C)$, si può ottenere immediatamente il valore del *fasore* $\overset{\circ}{Y}$ della tensione sinusoidale incognita $y(t)$:

$$\overset{\circ}{Y} = \frac{z_C}{R + z_C} \overset{\circ}{X} = \frac{1/(j2\pi f_0 C)}{R + 1/(j2\pi f_0 C)} \overset{\circ}{X} = \frac{1}{1 + j2\pi f_0 RC} a e^{j\theta} \tag{2.1.3}$$

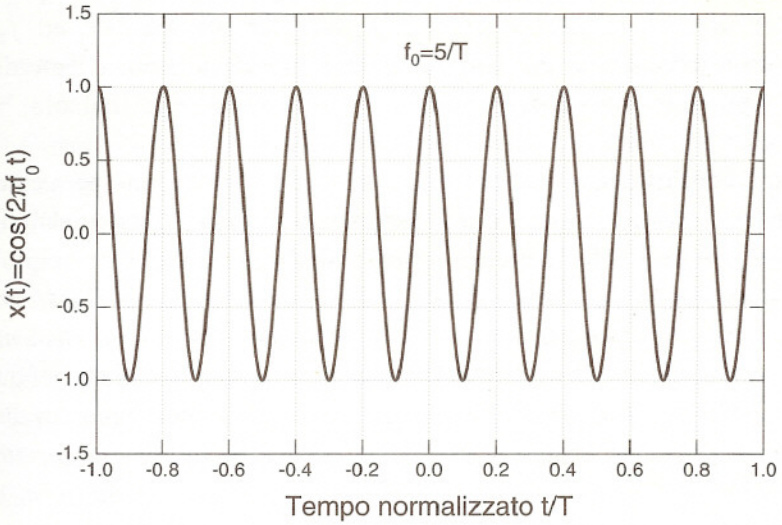
e da questo risalire poi all'espressione cercata di $y(t)$:

$$y(t) = |\overset{\circ}{Y}| \cdot \cos(2\pi f_0 t + \angle \overset{\circ}{Y}) \tag{2.1.4}$$

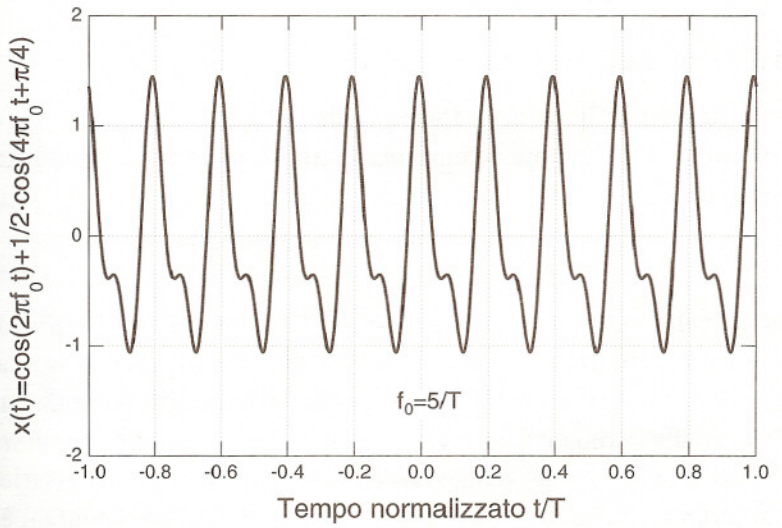
Il metodo fasoriale, di applicazione estremamente semplice, continua a rivelarsi utile anche quando il segnale di eccitazione $x(t)$ è come in Figura 2.2b, costituito cioè dalla somma di *più componenti sinusoidali* a diverse frequenze:

$$x(t) = a_1 \cos(2\pi f_1 t + \theta_1) + a_2 \cos(2\pi f_2 t + \theta_2) \tag{2.1.5}$$

In tal caso, stante la *linearità* del circuito è possibile applicare il principio di sovrapposizione degli effetti come sequenza di tre passi fondamentali:



(a)



(b)

Figura 2.2 (a) Segnale sinusoidale con $a = 1$, $f_0 = 5/T$ e $\theta = 0$; (b) segnale composito con $a_1 = 1$, $f_1 = f_0$, $\theta_1 = 0$ e $a_2 = 1/2$, $f_2 = 2f_0$, $\theta_2 = \pi/4$

i) si scompone il segnale d'ingresso $x(t)$ nelle due componenti elementari $x_1(t) = a_1 \cos(2\pi f_1 t + \theta_1)$ e $x_2(t) = a_2 \cos(2\pi f_2 t + \theta_2)$ con fasori $\overset{\circ}{X}_1 = a_1 \exp(j\theta_1)$ e $\overset{\circ}{X}_2 = a_2 \exp(j\theta_2)$ riferiti rispettivamente alle due frequenze f_1 ed f_2; ii) si determinano separatamente i due fasori $\overset{\circ}{Y}_1$ e $\overset{\circ}{Y}_2$ come se in ciascuno dei due casi agisse *singolarmente* la componente rispettivamente $x_1(t)$ e $x_2(t)$; iii) si

ricostruisce la forma d'onda globale $y(t)$ ricomponendo le due risposte parziali $y_1(t)$ e $y_2(t)$ ottenute dai due fasori $\overset{\circ}{Y}_1$ e $\overset{\circ}{Y}_2$ alle due frequenze f_1 ed f_2. Il procedimento è ovviamente più complicato rispetto al caso precedente della singola componente sinusoidale, ma solo per questioni di calcolo, non comportando alcuna ulteriore difficoltà concettuale.

Esaminando criticamente la questione, i presupposti che hanno permesso di risolvere anche il caso più complesso del segnale di Figura 2.2b sono due: i) la linearità del circuito, che ha consentito l'applicazione del principio di sovrapposizione degli effetti, ma soprattutto ii) *la possibilità di rappresentare il segnale come somma di un certo numero di componenti sinusoidali elementari* (2.1.5). Questa osservazione fa da introduzione al prossimo paragrafo nel quale si darà una risposta all'ulteriore quesito che sorge spontaneo generalizzando ulteriormente il problema appena esaminato: *Qual è il modo più appropriato di procedere quando il segnale $x(t)$ è periodico con andamento arbitrario, e in particolare non è sinusoidale?*

2.2 Analisi armonica dei segnali periodici

2.2.1 Sviluppo in serie di Fourier in forma reale polare

Come già affermato nel Capitolo 1, un segnale $x(t)$ è periodico se soddisfa la seguente relazione:

$$x(t) = x(t + T_0) \tag{2.2.1}$$

per ogni valore della variabile t. La grandezza T_0 rappresenta il *periodo* del segnale che è legato alla *frequenza di ripetizione* f_0 del segnale stesso dalla relazione $f_0 = 1/T_0$. *L'energia* E_x del segnale periodico è infinita, come discende immediatamente dalla definizione (1.3.1); in generale invece $x(t)$ ha *potenza* P_x finita, per calcolare la quale, come si può intuire, non è necessario il procedimento al limite definito dalla (1.3.4). Si dimostra facilmente infatti che, una volta eseguito il limite, si giunge al risultato più semplice

$$P_x = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} |x(t)|^2 \, dt \tag{2.2.2}$$

Analogamente, l'espressione del *valor medio* (1.3.7*a*) si semplifica, per il segnale periodico $x(t)$, come segue:

$$x_m = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\,dt \tag{2.2.3}$$

Ciò premesso, ci poniamo ancora una volta la domanda con cui si è chiuso il Paragrafo 2.1, e cioè: come trattare segnali periodici arbitrari, in particolare non sinusoidali? La risposta a questo quesito sta nella cosiddetta *analisi di Fourier* che costituisce la base della moderna teoria dei segnali. Infatti, sotto ipotesi piuttosto ampie, che in seguito elencheremo, un segnale reale periodico qualunque può essere espresso come *somma di oscillazioni sinusoidali di ampiezza, frequenza e fase opportune*, cioè in una forma che richiama la (2.1.5):

$$x(t) = A_0 + A_1 \cos(2\pi f_1 t + \vartheta_1) + A_2 \cos(2\pi f_2 t + \vartheta_2) + \ldots. \tag{2.2.4}$$

In particolare, le frequenze di oscillazione includono in generale la "frequenza zero" relativa al termine costante a_0, e sono *multiple intere della frequenza fondamentale* f_0, cosicché la (2.2.4) diventa

$$x(t) = A_0 + 2\sum_{k=1}^{\infty} A_k \cos(2\pi k f_0 t + \vartheta_k) \tag{2.2.5}$$

SVILUPPO IN SERIE DI FOURIER IN FORMA REALE POLARE

Questa rappresentazione del segnale prende il nome di *sviluppo in serie di Fourier;* più precisamente la relazione (2.2.5) costituisce l'*espressione in forma polare* dello sviluppo in serie di Fourier. Essa permette dunque di rappresentare un segnale reale $x(t)$ come somma di una costante A_0 e di una *serie* il cui k-esimo termine, detto k-esima *oscillazione armonica* (o *armonica* tout-court), ha ampiezza $A_k > 0$, frequenza kf_0 (la k-esima frequenza armonica) e fase iniziale ϑ_k.

Evidentemente, ogni particolare segnale $x(t)$ sarà caratterizzato da particolari insiemi di valori di A_k e ϑ_k. Dovremo quindi ricavare delle formule utili per il calcolo delle ampiezze e delle fasi delle varie armoniche e indicare condizioni matematiche che garantiscano la convergenza della serie (2.2.5). Il primo di questi problemi fu risolto dal matematico L. Eulero attorno alla fine del 1700 in connessione con lo studio delle corde vibranti, e fu ripreso alcuni anni più tardi da J.B. Fourier. Questi fu il primo a intuire l'importanza e la potenza della rappresentazione (2.2.5), che usò per risolvere questioni di trasmissione del calore. La *convergenza* della (2.2.5) fu dimostrata in seguito in maniera rigorosa da P.D. Dirichlet.

2.2.2 Sviluppo in serie di Fourier in forma complessa

Per semplificare gli sviluppi analitici si preferisce usare una forma alternativa della serie di Fourier. Richiamando le formule di Eulero delle funzioni trigonometriche

$$\cos(x)=\frac{e^{jx}+e^{-jx}}{2} \quad , \quad \sin(x)=\frac{e^{jx}-e^{-jx}}{2j}, \tag{2.2.6}$$

la (2.2.5) può essere riscritta come segue:

$$\begin{aligned} x(t) &= A_0 + 2\sum_{k=1}^{\infty} A_k \frac{e^{j(2\pi k f_0 t+\vartheta_k)}+e^{-j(2\pi k f_0 t+\vartheta_k)}}{2} = \\ &= A_0 + \sum_{k=1}^{\infty} A_k\, e^{j\vartheta_k} e^{j2\pi k f_0 t} + \sum_{k=1}^{\infty} A_k\, e^{-j\vartheta_k} e^{-j2\pi k f_0 t} \\ &= A_0 + \sum_{k=1}^{\infty} A_k\, e^{j\vartheta_k} e^{j2\pi k f_0 t} + \sum_{k=-\infty}^{-1} A_{-k}\, e^{-j\vartheta_{-k}} e^{j2\pi k f_0 t} \end{aligned} \tag{2.2.7}$$

Definiamo ora le quantità

$$X_0 \triangleq A_0 \quad , \quad X_k \triangleq A_k\, e^{j\vartheta_k},\ k=1,2,\ldots \quad , \quad X_k \triangleq A_{-k}\, e^{-j\vartheta_{-k}},\ k=\ldots,-2,-1 \tag{2.2.8}$$

Se si effettuano le opportune sostituzioni nella (2.2.7) si ricava

$$x(t) = X_0 + \sum_{k=1}^{\infty} X_k\, e^{j2\pi k f_0 t} + \sum_{k=-\infty}^{-1} X_k\, e^{j2\pi k f_0 t} = \sum_{k=-\infty}^{\infty} X_k\, e^{j2\pi k f_0 t} \tag{2.2.9}$$

che rappresenta l'*espressione in forma complessa* della serie di Fourier[1], e che risulta la più conveniente dal punto di vista del calcolo.

Determiniamo ora una espressione per il calcolo del coefficiente X_n, ove n deve intendersi *fissato*. A tal fine moltiplichiamo entrambi i membri della (2.2.9) per il fattore $e^{-j2\pi n f_0 t}$ e integriamo il risultato in un intervallo di ampiezza pari al periodo T_0 del segnale:

$$\int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi n f_0 t}\, dt = \int_{-T_0/2}^{T_0/2} \sum_{k=-\infty}^{+\infty} X_k\, e^{j2\pi k f_0 t} e^{-j2\pi n f_0 t}\, dt \tag{2.2.10}$$

Supponendo che la serie a secondo membro converga *uniformemente* (cosa che

---

[1] Tale rappresentazione può essere estesa nella stessa forma anche al caso di segnale $x(t)$ *complesso*.

peraltro non è stata dimostrata fino a questo momento), possiamo integrare *termine a termine*:

$$\int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi n f_0 t}\, dt = \sum_{k=-\infty}^{+\infty} X_k \int_{-T_0/2}^{T_0/2} e^{j2\pi(k-n)f_0 t}\, dt \tag{2.2.11}$$

Procediamo adesso con il calcolo dell'integrale a secondo membro della (2.2.11). Ricordando che $f_0 \cdot T_0 = 1$ si ha

$$\int_{-T_0/2}^{T_0/2} e^{j2\pi(k-n)f_0 t}\, dt = \frac{\exp[j2\pi(k-n)f_0 t]}{j2\pi(k-n)f_0}\Bigg|_{-T_0/2}^{T_0/2} =$$

$$= \frac{\exp[j\pi(k-n)] - \exp[-j\pi(k-n)]}{j2\pi(k-n)f_0} = \frac{\sin[\pi(k-n)]}{\pi(k-n)f_0} \tag{2.2.12}$$

Il valore dell'integrale è pertanto nullo se $k \neq n$, essendo $\sin[\pi(k-n)] = 0$. Se $k = n$ il risultato finale della (2.2.12) perde di significato, ma ponendo $k = n$ direttamente nell'espressione di partenza si ricava immediatamente che l'integrale cercato vale in questo caso T_0. Riassumendo:

$$\int_{-T_0/2}^{T_0/2} e^{j2\pi(k-n)f_0 t}\, dt = \begin{cases} T_0 & k = n \\ 0 & k \neq n \end{cases} \tag{2.2.13}$$

e sostituendo nella (2.2.11) il risultato ottenuto, si ha:

$$\int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi n f_0 t}\, dt = X_n\, T_0 \tag{2.2.14}$$

dalla quale si deduce infine l'espressione cercata del coefficiente X_k:

$$X_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi k f_0 t}\, dt \tag{2.2.15}$$

Questa relazione permette quindi di effettuare il calcolo dei *coefficienti della serie di Fourier* di un segnale $x(t)$ dato. In particolare, per $k = 0$ si ha

$$X_0 = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, dt \tag{2.2.16}$$

che coincide con l'espressione del valor medio del segnale.

2.2.3 Sviluppo in serie di Fourier in forma reale rettangolare

Abbiamo dunque ricavato due possibili espressioni per la serie di Fourier, e precisamente quella in forma polare (2.2.5) e quella in forma complessa (2.2.9); ne esiste anche una terza, detta *espressione in forma rettangolare*, che ricaviamo di seguito. Sviluppando le funzioni cosinusoidali della (2.2.5) si ha

$$x(t) = A_0 + 2\sum_{k=1}^{\infty} A_k \cos(2\pi k f_0 t + \vartheta_k) =$$

$$= A_0 + 2\sum_{n=1}^{\infty} A_k \left[\cos(2\pi k f_0 t)\cos\vartheta_k - \sin(2\pi k f_0 t)\sin\vartheta_k\right] \qquad (2.2.17)$$

Se adesso si definiscono le quantità $a_0 \triangleq A_0$, $a_k \triangleq A_k \cos\vartheta_k$ e $b_k \triangleq A_k \sin\vartheta_k$, con $k = 1,2,\ldots$ si ricava la relazione cercata:

$$x(t) = a_0 + 2\sum_{k=1}^{\infty}\left[a_k \cos(2\pi k f_0 t) - b_k \sin(2\pi k f_0 t)\right] \qquad (2.2.18)$$

Il lettore dimostri che i coefficienti dell'espressione in forma rettangolare $\{a_k, b_k\}$ sono legati a quelli relativi all'espansione in forma complessa $\{X_k\}$ dalle relazioni

$$a_k = \Re[X_k] = \frac{1}{T_0}\int_{[T_0]} x(t)\cos(2\pi k f_0 t)\,dt \qquad (2.2.19)$$

$$b_k = \Im[X_k] = -\frac{1}{T_0}\int_{[T_0]} x(t)\sin(2\pi k f_0 t)\,dt \qquad (2.2.20)$$

Nelle equazioni precedenti, la notazione $\int_{[T_0]}$ sta a indicare che l'integrale può essere esteso a un qualunque intervallo temporale di ampiezza T_0. Per ragioni di simmetria, è buona norma scegliere l'intervallo $[-T_0/2, T_0/2]$.

2.3 Il criterio di Dirichlet

Ricordiamo che negli sviluppi analitici appena visti, e precisamente nel passaggio dalla (2.2.10) alla (2.2.11), è stata ipotizzata la convergenza uniforme della serie (2.2.9). Per i segnali che si incontrano comunemente nelle applicazioni pratiche, questa ipotesi è sempre verificata; spesso però, per schematizzare fenomeni fisici, si fa ricorso a funzioni che non rappresentano esattamente i segnali in esame, ma che offrono il vantaggio non indifferente di una maggiore *semplicità*. Per tali funzioni, tuttavia, non è più assicurata in

generale la possibilità di uno sviluppo in serie di Fourier e diventa quindi necessario disporre di criteri che garantiscano la correttezza di tale sviluppo.

Consideriamo ad esempio il *segnale a dente di sega* rappresentato in Figura 2.3a; dal punto di vista matematico, esso presenta all'istante T_0 una discontinuità di prima specie (non eliminabile), in corrispondenza della quale esistono finiti e diversi tra loro i limiti destro e sinistro del segnale: $x(T_0^+) \neq x(T_0^-)$. Ovviamente, non avremo mai nella realtà un fenomeno fisico che si manifesta con tale andamento a causa dell'impossibilità di riscontrare una discontinuità nel segnale. Tuttavia, avremo a che fare con segnali che possono essere *ben approssimati* da un andamento discontinuo, come quello mostrato in Figura 2.3b, che è tipico dei circuiti di pilotaggio dei tubi catodici degli apparecchi televisivi.

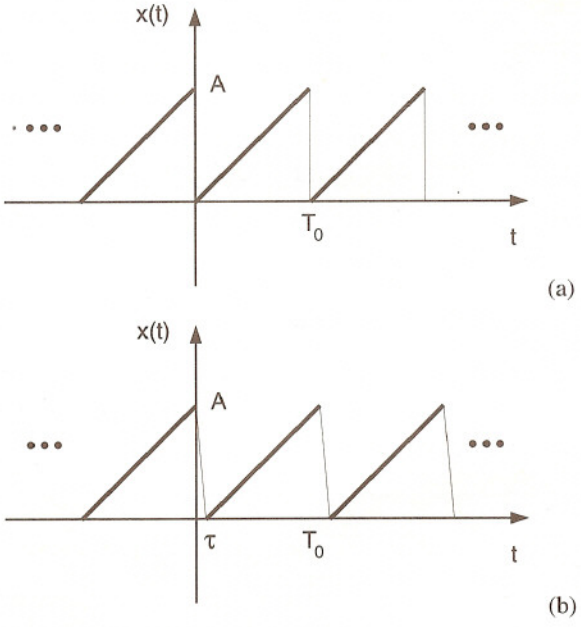


Figura 2.3 Segnali a dente di sega di ampiezza A: (a) ideale; (b) reale, con $\tau << T_0$

È quindi lecito domandarsi se per il segnale a dente di sega idealizzato, e per altri che hanno andamenti di particolare utilità e ai quali si applicano considerazioni analoghe, sia corretto utilizzare la rappresentazione (2.2.9) con i coefficienti espressi dalla (2.2.15).

Un insieme di condizioni sufficienti che garantiscono la possibilità di sviluppare un segnale in serie di Fourier è il cosiddetto *criterio di Dirichlet* che può essere enunciato come segue:

- *se* $x(t)$ è assolutamente integrabile sul periodo T_0 (cioè se verifica la condizione $\int_{-T_0/2}^{T_0/2} |x(t)|\,dt < \infty$);
- *se* $x(t)$ è continua o presenta in un periodo un numero finito di discontinuità di prima specie;
- *se* $x(t)$ è derivabile rispetto al tempo nel periodo, escluso al più un numero finito di punti nei quali esistono finite la derivata destra e sinistra,

allora la serie di Fourier converge al valore assunto dalla funzione $x(t)$ *nei punti in cui questa è continua, e alla semisomma dei limiti destro e sinistro nei punti in cui* $x(t)$ *presenta le eventuali discontinuità di prima specie.*

La terza ipotesi del criterio può anche essere sostituita con la seguente, che risulta del tutto equivalente:

- *se* il segnale presenta un numero finito di massimi e minimi nel periodo

Alla luce di questo criterio è possibile adesso affermare con sicurezza che anche la funzione a dente di sega di Figura 2.3a può essere sviluppata in serie di Fourier. Nel punto di discontinuità il valore cui la serie converge è pari a $x(T_0) = [x(T_0^+) + x(T_0^-)]/2 = A/2$.

2.4 Spettri di ampiezza e di fase

Dunque, ogni segnale $x(t)$ che soddisfi il criterio di Dirichlet può essere rappresentato con lo sviluppo in serie di Fourier (2.2.9) ove i coefficienti X_k sono dati dalla (2.2.15). Ripetiamo per completezza qui di seguito queste due relazioni:

$$x(t) = \sum_{k=-\infty}^{\infty} X_k e^{j2\pi k f_0 t} \quad , \quad X_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) e^{-j2\pi k f_0 t} dt$$

La seconda delle due è una equazione di *analisi* che permette di stabilire qual è il contenuto in termini di oscillazioni armoniche del segnale (in una parola, di *analizzare* il segnale). La prima delle due, viceversa, è una equazione di *sintesi* che, note le ampiezze e fasi delle varie armoniche (cioè noti i coefficienti di Fourier) permette di ricostruire, cioè *sintetizzare*, il segnale dato a partire dalle proprie componenti frequenziali (armoniche). Evidentemente, l'equazione di sintesi prevede l'uso di *infinite* armoniche per ricostruire il segnale. D'altronde,

condizione necessaria alla convergenza della serie è che l'ampiezza $|X_k|$ delle armoniche *tenda a zero* quando $k \to \infty$. Questo comporta che le armoniche più "importanti" ai fini della sintesi del segnale sono in numero limitato, e che quindi la serie può essere sostituita ai fini pratici con una sommatoria di un numero finito di termini (come si discuterà in dettaglio nel Paragrafo 2.7).

Le equazioni di analisi e sintesi permettono dunque di stabilire una corrispondenza tra il segnale $x(t)$ e la *sequenza* X_k costituita dai coefficienti della serie (coefficienti di Fourier o di Eulero). Indicheremo tale corrispondenza con la seguente scrittura:

$$x(t) \Leftrightarrow X_k \tag{2.4.1}$$

Questo tipo di notazione suggerisce che la conoscenza dell'andamento del segnale $x(t)$ in ambito temporale è di fatto *equivalente* alla conoscenza della *successione* dei coefficienti di Fourier X_k in ambito frequenziale, nel senso che il passaggio dall'un dominio all'altro è immediato attraverso le relazioni di analisi e sintesi (2.2.15) e (2.2.9). Naturalmente, la sequenza X_k è in generale *complessa*; per rappresentarla è conveniente tracciare due grafici che prendono il nome di *spettro di ampiezza* e *spettro di fase*[2]. Il primo illustra l'andamento dell'ampiezza (modulo) dei coefficienti X_k, il secondo ne illustra l'andamento della fase, entrambi in funzione dell'ordine k del coefficiente o del valore della k-esima frequenza armonica kf_0. Esempi stilizzati di queste rappresentazioni sono riportati in Figura 2.4 e Figura 2.5.

Gli spettri di ampiezza e di fase del segnale sono *a righe*, cioè *discreti*, in quanto sono definiti solo in corrispondenza delle frequenze armoniche, che formano appunto una successione discreta. La rappresentazione degli spettri come "righe" di ampiezza proporzionale all'ampiezza o alla fase delle componenti armoniche ha un'origine ben precisa. Gli spettri di ampiezza a righe dei segnali periodici vengono infatti misurati mediante strumenti elettronici chiamati *analizzatori di spettro*, sullo schermo dei quali si ottiene una rappresentazione molto simile alla Figura 2.4. Gli analizzatori rappresentano gli spettri di ampiezza solo per valori *positivi* delle frequenze. Questo è giustificato dalle proprietà di simmetria degli spettri discusse nel paragrafo a seguire.

---

[2] Il termine "spettro" deve intendersi nel significato di "rappresentazione, visione" e nasce in fisica nel campo della spettroscopia in cui si analizza la composizione dei materiali attraverso le "righe" di emissione caratteristiche dei diversi elementi chimici.

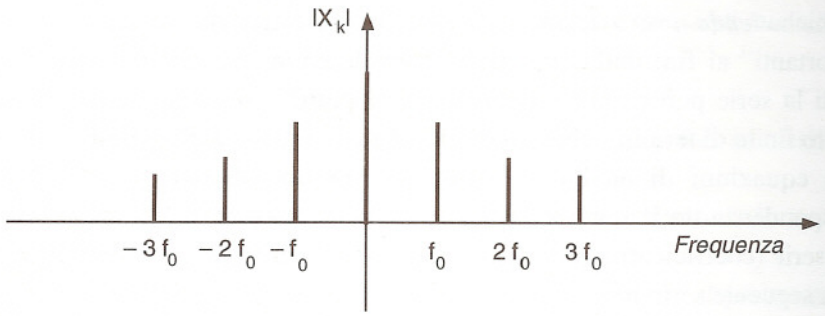


Figura 2.4 Spettro di ampiezza

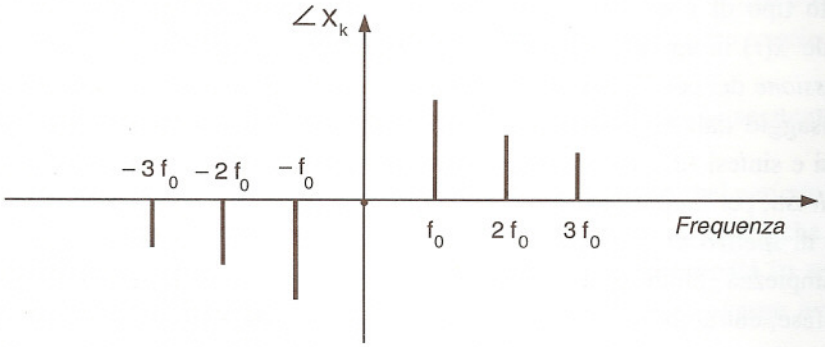


Figura 2.5 Spettro di fase

Esempio 2.1

Consideriamo il segnale

$$x(t) = a\cos(2\pi f_0 t) \tag{E2.1.1}$$

Esso rappresenta un'oscillazione cosinusoidale di frequenza f_0; il periodo del segnale è $T_0 = 1/f_0$. Se si confronta l'espressione in forma polare della serie di Fourier di un segnale generico

$$x(t) = A_0 + 2\sum_{k=1}^{\infty} A_k \cos(2\pi k f_0 t + \vartheta_k) \tag{E2.1.2}$$

con la (E2.1.1) si ricava che

$$A_1 = \frac{a}{2},\ \vartheta_1 = 0; \quad A_k, \vartheta_k = 0\ \ \forall k \neq 1 \tag{E2.1.3}$$

ovvero

$$X_1 = \frac{a}{2},\ X_{-1} = \frac{a}{2};\quad X_k = 0\quad \forall k \neq \pm 1 \tag{E2.1.4}$$

e gli spettri del segnale sono quelli di Figura 2.6*a*-2.6*b*.

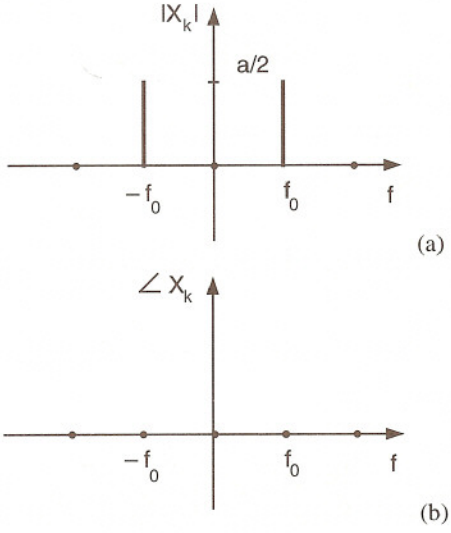


Figura 2.6 Spettri di ampiezza (a) e fase (b) del segnale dell'Esempio 2.1

❑

Esempio 2.2

Consideriamo il segnale

$$x(t) = a\sin(2\pi f_0 t) = a\cos\left(2\pi f_0 t - \frac{\pi}{2}\right) \tag{E2.2.1}$$

Tenendo conto del procedimento usato nell'Esempio 2.1, si trova che

$$A_1 = \frac{a}{2},\ \vartheta_1 = -\frac{\pi}{2};\quad A_k, \vartheta_k = 0\ \ \forall k \neq 1 \tag{E2.2.2}$$

ovvero

$$X_1 = \frac{a}{2}e^{-j\frac{\pi}{2}},\ X_{-1} = \frac{a}{2}e^{j\frac{\pi}{2}};\quad X_k = 0\ \ \forall k \neq \pm 1 \tag{E2.2.3}$$

Lo spettro di ampiezza è uguale a quello dell'esempio precedente, mentre lo spettro di fase è mostrato in Figura 2.7.

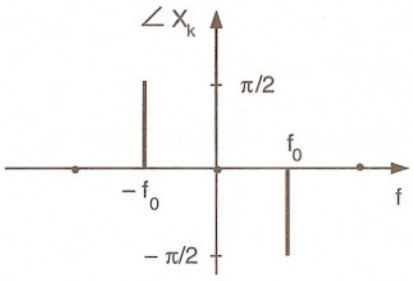


Figura 2.7 Spettro di fase del segnale dell'Esempio 2.2

❑

2.5 Proprietà dello spettro di un segnale reale periodico

Simmetria

Ricordando che il nostro punto di partenza è stata la rappresentazione (2.2.5) di un segnale *reale*, osserviamo che

$$X_0 = A_0, \quad X_k = A_k\, e^{j\vartheta_k}, \quad X_{-k} = A_k\, e^{-j\vartheta_k}; \quad k = 1,2,\ldots \tag{2.5.1}$$

e che quindi i coefficienti X_k dello sviluppo in serie di Fourier in forma complessa di un segnale reale godono della proprietà di *simmetria coniugata o Hermitiana*, ovvero

$$X_{-k} = X_k^* \Leftrightarrow \begin{cases} |X_k| = |X_{-k}| \\ \angle X_k = -\angle X_{-k} \end{cases} \tag{2.5.2}$$

Linearità

Consideriamo ora due segnali $x(t)$ e $y(t)$, entrambi periodici dello stesso periodo T_0 e aventi come coefficienti di Fourier rispettivamente X_k e Y_k. Il segnale

$$z(t) = a\, x(t) + b\, y(t) \tag{2.5.3}$$

dato dalla combinazione lineare di $x(t)$ e $y(t)$ è periodico di periodo T_0, e il coefficiente k-esimo Z_k della sua serie di Fourier è

$$Z_k = a\, X_k + b\, Y_k \tag{2.5.4}$$

Tale proprietà di *linearità* dei coefficienti di Fourier deriva direttamente dalla medesima proprietà dell'integrale; si ha infatti che

$$Z_k = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} z(t)\, e^{-j2\pi k f_0 t}\, dt = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} [a\, x(t) + b\, y(t)]\, e^{-j2\pi k f_0 t}\, dt =$$

$$= \frac{a}{T_0}\int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi k f_0 t}\, dt + \frac{b}{T_0}\int_{-T_0/2}^{T_0/2} y(t)\, e^{-j2\pi k f_0 t}\, dt = a\, X_k + b\, Y_k \qquad (2.5.5)$$

Naturalmente, lo sviluppo in serie di $z(t)$ è costituito da una somma di oscillazioni aventi *le stesse frequenze* di quelle che compongono i segnali $x(t)$ e $y(t)$; pertanto, in generale una combinazione lineare di segnali aventi il medesimo periodo T_0 non introduce nuove armoniche.

Potremmo elencare e dimostrare molte altre proprietà dei coefficienti di Fourier (ad esempio, il comportamento degli X_k nei confronti di una traslazione del segnale, o di una operazione di derivata temporale, o di integrale, ecc.). Le più semplici vengono proposte come esercizi in calce al capitolo, che possono essere facilmente risolti dal lettore; queste proprietà verranno comunque discusse in maggior dettaglio e in una forma sostanzialmente equivalente nel Capitolo 3, a proposito della *trasformata di Fourier* per segnali non periodici, al quale si rimanda.

Esempio 2.3

Consideriamo il segnale detto *treno di impulsi rettangolari* di *durata* T e *periodo* T_0 ($T < T_0$) rappresentato in Figura 2.8. Per questo segnale si definisce il parametro *duty-factor* (o *duty-cycle*) $\delta = T/T_0$ che esprime il rapporto tra la durata T di ciascun impulso e il periodo di ripetizione del segnale T_0.

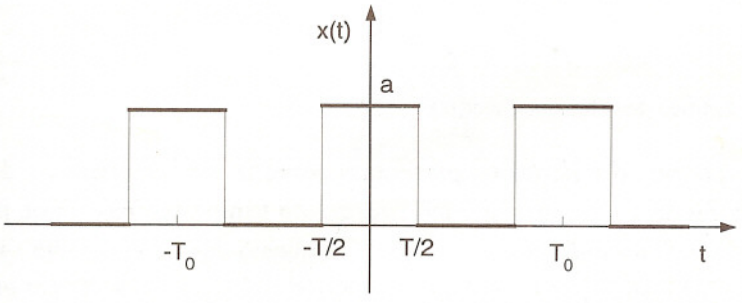


Figura 2.8 Treno di impulsi rettangolari

Per rappresentare più comodamente il treno di impulsi rettangolari è utile definire la seguente funzione:

$$\text{rect}(\alpha) \triangleq \begin{cases} 1 & |\alpha| < 1/2 \\ 1/2 & |\alpha| = 1/2 \\ 0 & \text{altrove} \end{cases} \qquad \text{(E2.3.1)}$$

il cui andamento (impulso rettangolare) è rappresentato in Figura 2.9. La discussione relativa ai punti di discontinuità del segnale a dente di sega di Figura 2.3 è valida anche per il segnale rect(·). Ancora una volta, questa semplice funzione rappresenta un'astrazione matematica utile per schematizzare impulsi che hanno un *tempo di salita* molto breve rispetto alla propria *durata*. Utilizzando questa funzione è possibile scrivere la seguente espressione per il treno di impulsi rettangolari:

$$x(t) = \sum_{n=-\infty}^{+\infty} a \cdot \text{rect}\left(\frac{t - n\,T_0}{T}\right) \qquad \text{(E2.3.2)}$$

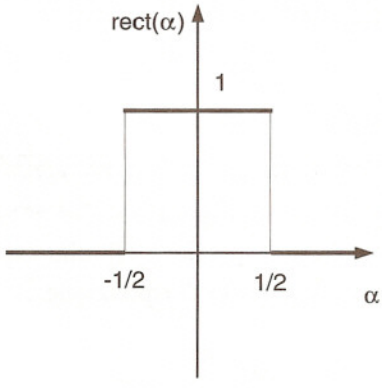


Figura 2.9 Grafico della funzione rect(α)

Il segnale periodico è infatti rappresentato come la *sovrapposizione* di infiniti impulsi di durata T ottenuti ciascuno *ritardando* l'impulso "base" non periodico $\text{rect}(t/T)$ di nT_0 secondi con $n = 0, \pm 1, \ldots$ In questo caso diremo che l'impulso-base $\text{rect}(t/T)$ è stato *periodicizzato* con periodo di ripetizione T_0 per ottenere il segnale $x(t)$. Il lettore può dimostrare facilmente che il segnale $x(t)$ nella forma di segnale-base ripetuto (E2.3.2) è effettivamente periodico di periodo T_0.

Calcoliamo ora i coefficienti dello sviluppo in serie del segnale in esame utilizzando la (2.2.15):

$$X_k = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} a\cdot \mathrm{rect}\left(\frac{t}{T}\right) e^{-j2\pi k f_0 t}\,dt = \frac{1}{T_0}\int_{-T/2}^{T/2} a\, e^{-j2\pi k f_0 t}\,dt = \frac{a}{T_0}\frac{e^{-j2\pi k f_0 t}}{(-j2\pi k f_0)}\Bigg|_{-T/2}^{T/2} =$$

$$= \frac{a}{T_0}\frac{e^{-j\pi k f_0 T} - e^{j\pi k f_0 T}}{(-j2\pi k f_0)} = \frac{a}{T_0}\frac{\sin(\pi k f_0 T)}{\frac{\pi k}{T_0}} = \frac{aT}{T_0}\frac{\sin\left(\frac{\pi k T}{T_0}\right)}{\frac{\pi k T}{T_0}} \qquad \text{(E2.3.3)}$$

Per esprimere X_k in una forma più concisa, definiamo una ulteriore funzione notevole:

$$\mathrm{sinc}(\alpha) \triangleq \frac{\sin(\pi\alpha)}{\pi\alpha}, \qquad \text{(E2.3.4)}$$

per cui

$$X_k = \frac{aT}{T_0}\mathrm{sinc}\left(\frac{kT}{T_0}\right) = a\,\delta\,\mathrm{sinc}(k\delta) \qquad \text{(E2.3.5)}$$

La funzione sinc(·), il cui grafico è mostrato in Figura 2.10, verrà usata molto di frequente nella pagine a seguire, e quindi è importante evidenziarne alcune caratteristiche peculiari.

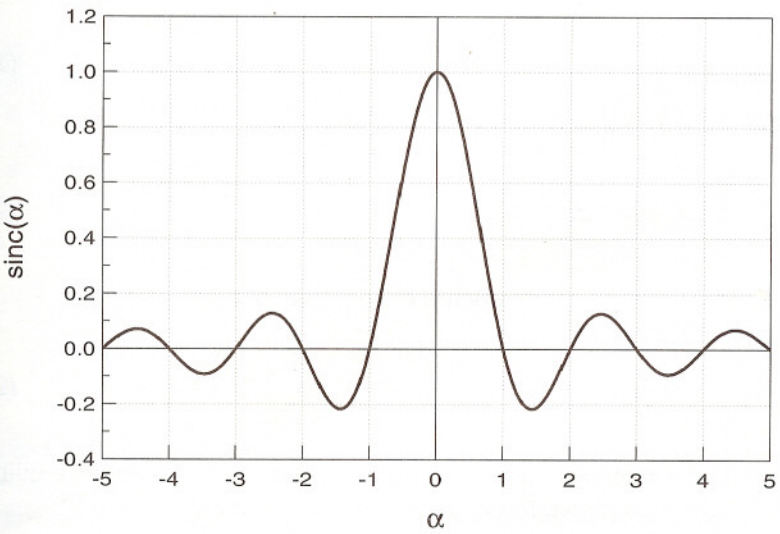


Figura 2.10 Andamento della funzione sinc(α)

Come si nota dalla Figura 2.10, essa si annulla per tutti i valori interi del suo argomento α diversi da zero, mentre nell'origine assume valore unitario. La discontinuità in questo punto è stata infatti eliminata, osservando che

$$\lim_{\alpha \to 0} \frac{\sin(\pi\alpha)}{\pi\alpha} = 1 \qquad \text{(E2.3.6)}$$

Inoltre, $\text{sinc}(\alpha)$ è una funzione pari in quanto rapporto di due funzioni dispari, tende ad annullarsi al tendere di α all'infinito, e in particolare soddisfa la relazione $\text{sinc}(\alpha) \leq 1/(\pi\,|\,\alpha\,|)$. Il calcolo dei punti di massimo e minimo della funzione porta invece a una equazione *trascendente* (che il lettore può facilmente ricavare) e che non verrà presa in considerazione per brevità.

Scegliamo ora nel nostro treno di impulsi i valori particolari $a = 1$ e $\delta = 0.5$. Gli spettri di ampiezza e fase del segnale hanno allora l'andamento illustrato nelle Figure 2.11-2.12. Si nota che lo spettro di ampiezza è *simmetrico* rispetto alla frequenza 0, mentre lo spettro di fase assume soltanto i valori 0 e $\pm\pi$ (il coefficiente di Fourier è infatti *reale*), e conserva simmetria *dispari*.

2.6 Segnali pari, dispari e alternativi

Il segnale (periodico) $x(t)$ è *pari* se verifica la relazione $x(t) = x(-t)$. In tal caso, il generico coefficiente X_k della serie di Fourier di $x(t)$ è una funzione *pari* di k, cioè

$$X_k = X_{-k} \qquad \text{(2.6.1)}$$

Infatti, dalla (2.2.15) è immediato verificare che

$$X_{-k} = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, e^{j2\pi k f_0 t}\, dt \qquad \text{(2.6.2)}$$

nella quale, effettuando il cambiamento di variabile $\alpha = -t$, si ottiene

$$X_{-k} = -\frac{1}{T_0} \int_{T_0/2}^{-T_0/2} x(-\alpha)\, e^{-j2\pi k f_0 \alpha}\, d\alpha = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(\alpha)\, e^{-j2\pi k f_0 \alpha}\, d\alpha = X_k \qquad \text{(2.6.3)}$$

Poiché in generale sappiamo che $X_k = \left(X_{-k}\right)^*$, segue che X_k è *reale*, e quindi è possibile riscrivere la serie di Fourier in forma semplificata:

$$x(t) = X_0 + \sum_{k=1}^{+\infty} X_k\, e^{j2\pi k f_0 t} + \sum_{k=-\infty}^{-1} X_k\, e^{j2\pi k f_0 t} \qquad \text{(2.6.4)}$$

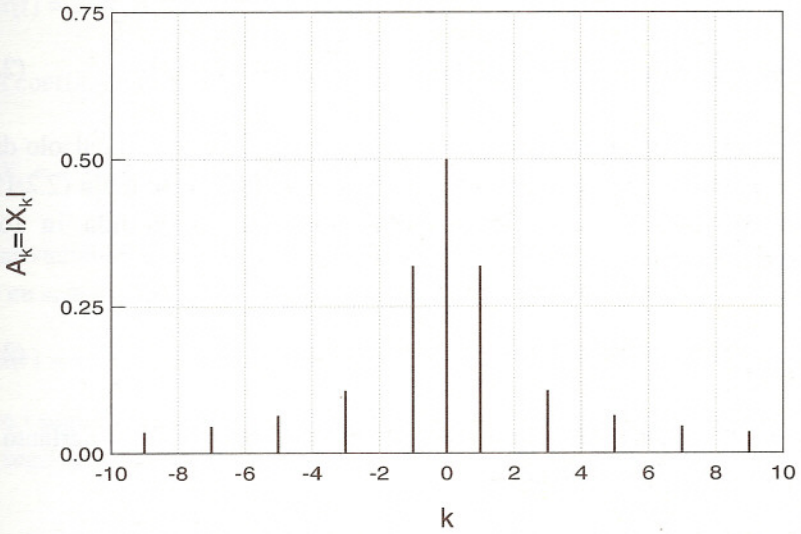


Figura 2.11 Spettro di ampiezza del segnale di Figura 2.8 per $\delta = 0.5$ e $a = 1$

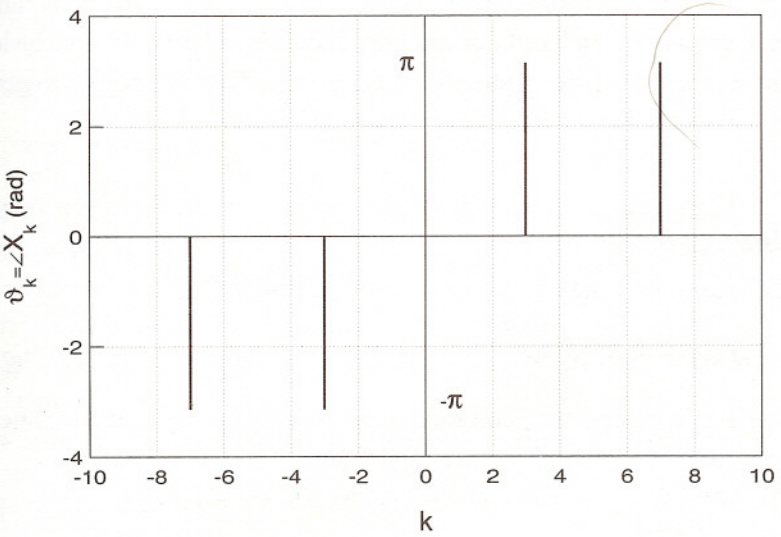


Figura 2.12 Spettro di fase del segnale di Figura 2.8 per $\delta = 0.5$ e $a = 1$

e, cambiando segno all'indice nella seconda sommatoria, si ottiene

$$x(t) = X_0 + \sum_{k=1}^{+\infty} X_k \, e^{j2\pi k f_0 t} + \sum_{k=1}^{+\infty} X_{-k} \, e^{-j2\pi k f_0 t} \qquad (2.6.5)$$

Poiché $X_k = X_{-k}$, unendo le due sommatorie si ha infine

$$x(t) = X_0 + \sum_{n=1}^{+\infty} X_k \left(e^{j2\pi k f_0 t} + e^{-j2\pi k f_0 t} \right) = X_0 + 2 \sum_{k=1}^{+\infty} X_k \cos(2\pi k f_0 t) \tag{2.6.6}$$

Il segnale è dunque esprimibile in *serie di soli coseni*. Inoltre, il calcolo di X_k può essere effettuato con una formula semplificata. Infatti, se nella (2.2.15) si riscrive l'esponenziale complesso nella funzione integranda in forma rettangolare si ottiene

$$X_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \cos(2\pi k f_0 t)\, dt - \frac{j}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \sin(2\pi k f_0 t)\, dt \tag{2.6.7}$$

La funzione integranda del primo integrale della (2.6.7) è pari e pertanto tale integrale può essere calcolato come

$$\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \cos(2\pi k f_0 t)\, dt = \frac{2}{T_0} \int_{0}^{T_0/2} x(t) \cos(2\pi k f_0 t)\, dt \tag{2.6.8}$$

Invece il contributo del secondo integrale della (2.6.7) è nullo poiché la funzione integranda è dispari e l'intervallo di integrazione è simmetrico. In conclusione si ricava che, se $x(t)$ è reale e pari, il k-esimo coefficiente della sua serie di Fourier può essere calcolato con la formula semplificata

$$X_k = \frac{2}{T_0} \int_{0}^{T_0/2} x(t) \cos(2\pi k f_0 t)\, dt \in \Re \tag{2.6.9}$$

Un segnale (periodico) $x(t)$ è *dispari* se verifica la relazione

$$x(t) = -x(-t) \tag{2.6.10}$$

Allora è possibile dimostrare, con considerazioni identiche a quelle fatte per i segnali pari, che la sua serie di Fourier gode delle proprietà seguenti:

- il coefficiente X_k della serie è una funzione *dispari* di k, cioè

$$X_{-k} = -X_k \tag{2.6.11}$$

  ed è quindi *immaginario puro*. Ne consegue che $X_0 = 0$;
- il segnale è sviluppabile in *serie di soli seni*:

$$x(t) = 2j\sum_{k=1}^{+\infty} X_k \sin(2\pi k f_0 t) \tag{2.6.12}$$

- il coefficiente X_k può essere calcolato con la seguente formula semplificata:

$$X_k = -\frac{2j}{T_0}\int_0^{T_0/2} x(t)\sin(2\pi k f_0 t)\,dt \tag{2.6.13}$$

Relativamente alle simmetrie *pari* e *dispari* di un segnale, è interessante notare che un segnale reale *arbitrario* $x(t)$ può essere sempre scomposto come

$$x(t) = x_p(t) + x_d(t) \tag{2.6.14}$$

dove i segnali $x_p(t)$ e $x_d(t)$ sono rispettivamente la *parte pari* e la *parte dispari* del segnale dato:

$$x_p(t) \triangleq \frac{x(t)+x(-t)}{2} \tag{2.6.15}$$

$$x_d(t) \triangleq \frac{x(t)-x(-t)}{2} \tag{2.6.16}$$

Sulla base di questa scomposizione si può scrivere

$$X_k = X_{pk} + X_{dk} \tag{2.6.17}$$

ove evidentemente X_{pk} (il k-esimo coefficiente di $x_p(t)$) è *reale*, mentre X_{dk} (k-esimo coefficiente di $x_d(t)$) è *immaginario puro*. Si conclude allora che la parte reale del coefficiente di Fourier di un segnale reale arbitrario $x(t)$ rappresenta il coefficiente dello sviluppo della sua parte pari $x_p(t)$, mentre la parte immaginaria di X_k è (a meno del fattore j) il coefficiente dello sviluppo della parte dispari $x_d(t)$ di $x(t)$.

Un segnale periodico $x(t)$ è *alternativo* se verifica la relazione

$$x(t + T_0/2) = -x(t) \tag{2.6.18}$$

cioè se l'andamento del segnale in un qualunque *semiperiodo* $t_0 \le t < t_0 + T_0/2$ è identico all'andamento nel semiperiodo precedente $t_0 - T_0/2 \le t < t_0$, cambiato di segno (si veda la Figura 2.13 per un esempio). Allora il coefficiente X_k della serie di Fourier di $x(t)$ è *nullo* per tutti i valori *pari* dell'indice k. Infatti X_k è dato da

$$X_k = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi k f_0 t}\, dt = \frac{1}{T_0}\int_{0}^{T_0/2} x(t)\, e^{-j2\pi k f_0 t}\, dt + \frac{1}{T_0}\int_{-T_0/2}^{0} x(t)\, e^{-j2\pi k f_0 t}\, dt \qquad (2.6.19)$$

avendo suddiviso l'intervallo di integrazione $\left[-T_0/2,\, T_0/2\right]$ in due semiperiodi.

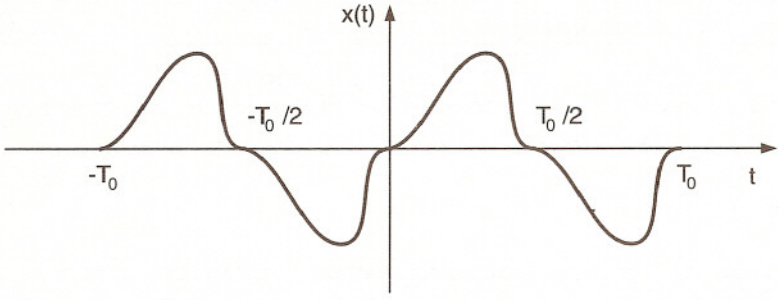


Figura 2.13 Segnale alternativo

Il secondo integrale a secondo membro della (2.6.19) può essere riscritto come

$$\frac{1}{T_0}\int_{T_0/2}^{T_0} x(t)\, e^{-j2\pi k f_0 t}\, dt = \frac{1}{T_0}\int_{0}^{T_0/2} x(\alpha + T_0/2)\, e^{-j2\pi k f_0 (\alpha + T_0/2)}\, d\alpha$$

$$= -\frac{e^{-j\pi k}}{T_0}\int_{0}^{T_0/2} x(\alpha)\, e^{-j2\pi k f_0 \alpha}\, d\alpha = -\frac{(-1)^k}{T_0}\int_{0}^{T_0/2} x(\alpha)\, e^{-j2\pi k f_0 \alpha}\, d\alpha \qquad (2.6.20)$$

Sostituendo la (2.6.20) nella (2.6.19) si ricava

$$X_k = \frac{1-(-1)^k}{T_0}\int_{0}^{T_0/2} x(t)\, e^{-j2\pi k f_0 t}\, dt \qquad (2.6.21)$$

da cui

$$X_k = \begin{cases} 0 & k \text{ pari} \\ \dfrac{2}{T_0}\displaystyle\int_{0}^{T_0/2} x(t)\, e^{-j2\pi k f_0 t}\, dt & k \text{ dispari} \end{cases} \qquad (2.6.22)$$

e l'espressione (2.2.9) della serie di Fourier si semplifica come segue:

$$x(t) = \sum_{\substack{k=-\infty \\ k \text{ dispari}}}^{+\infty} X_k\, e^{j2\pi k f_0 t} = \sum_{p=-\infty}^{+\infty} X_{2p+1}\, e^{j2\pi (2p+1) f_0 t} \qquad (2.6.23)$$

avendo effettuato il cambiamento di variabile $k = 2p+1$ per rappresentare l'insieme dei numeri relativi dispari.

Esempio 2.4

Calcoliamo lo sviluppo in serie di Fourier del segnale *onda quadra* $y(t)$ rappresentato in Figura 2.14*a*.

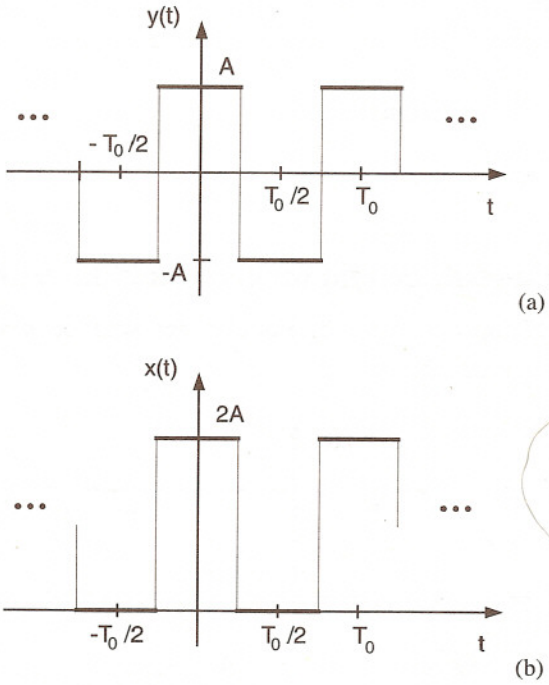


Figura 2.14 Segnale onda quadra (a) e treno di impulsi rettangolari (b)

Vediamo come esprimere questo segnale in funzione del treno di impulsi rettangolari $x(t)$ di Figura 2.8. Se per quest'ultimo assegnamo i valori $a = 2A$ e $T = T_0/2$ ($\delta = 1/2$), otteniamo il segnale $x(t)$ di Figura 2.14*b*. Allora è immediato vedere che

$$y(t) = x(t) - A \qquad \text{(E2.4.1)}$$

Il segnale costante pari ad A a secondo membro della (E2.4.1) può essere considerato periodico con sviluppo in serie di Fourier ridotto al solo termine costante e pari ovviamente ad a. Per la linearità della serie di Fourier si ha quindi

$$Y_k = \begin{cases} X_0 - A & k = 0 \\ X_k & k \neq 0 \end{cases} \tag{E2.4.2}$$

Poiché

$$X_k = 2A\delta \ \text{sinc}(\delta k) = A \ \text{sinc}\left(\frac{k}{2}\right) \tag{E2.4.3}$$

la (E2.4.2) diventa

$$Y_k = \begin{cases} 0 & k = 0 \\ A \ \text{sinc}(k/2) & \text{altrimenti} \end{cases} \tag{E2.4.4}$$

❑

2.7 Sintesi del segnale con un numero limitato di armoniche

Calcoliamo lo sviluppo in serie di Fourier del segnale *onda quadra* $x(t)$ rappresentato in Figura 2.15.

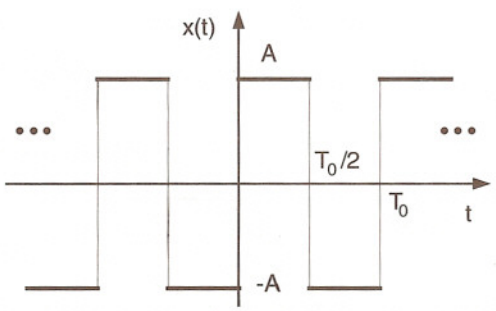


Figura 2.15 Segnale onda quadra

Il segnale è dispari e alternativo. Pertanto il calcolo del coefficiente X_k della sua serie di Fourier può essere effettuato con la formula semplificata (2.6.13) o con la (2.6.21). Scegliendo per semplicità la prima delle due, si ha:

$$X_k = -\frac{2j}{T_0}\int_0^{T_0/2} x(t)\sin(2\pi k f_0 t)\,dt = -j\frac{2A}{T_0}\int_0^{T_0/2}\sin(2\pi k f_0 t)\,dt$$
$$= j\frac{2A}{2\pi k f_0 T_0}\cos(2\pi k f_0 t)\Big|_{t=0}^{t=T_0/2} = j\frac{A}{\pi k}\left[\cos(\pi k) - 1\right] = j\frac{A}{\pi k}\left[(-1)^k - 1\right] \tag{2.7.1}$$

da cui

$$X_k = \begin{cases} 0 & k \text{ pari} \\ \dfrac{2A}{j\pi k} & k \text{ dispari} \end{cases} \tag{2.7.2}$$

Come ci aspettavamo, la sequenza $\{X_k\}$ dei coefficienti di Fourier dell'onda quadra (qui in versione *dispari)* è immaginaria pura e dispari, e i coefficienti di ordine pari sono tutti nulli (l'onda quadra è *alternativa*). Lo spettro risultante è quello di Figura 2.16.

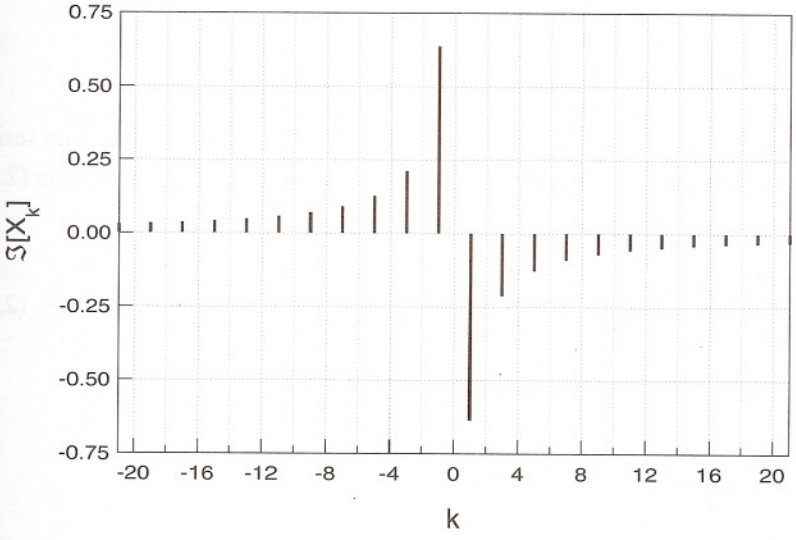


Figura 2.16 Spettro del segnale in Figura 2.15 con $A = 1$

Lo sviluppo in serie di Fourier di $x(t)$ (in soli seni) si ottiene sostituendo l'espressione degli X_k (2.7.2) nella (2.6.12):

$$x(t) = \frac{4A}{\pi} \sum_{k=1}^{+\infty} \frac{\sin(2\pi k f_0 t)}{k} \tag{2.7.3}$$

Si nota che l'ampiezza delle righe tende a zero come $1/k$ quando $k \to \infty$. Questa è la minima "velocità" con cui lo spettro di ampiezza può decrescere, affinché la serie di funzioni (2.7.3) risulti convergente. In altri termini, l'ampiezza delle componenti armoniche ad "alte frequenze" (cioè le armoniche di ordine elevato) è ancora relativamente grande rispetto a quella delle prime armoniche.

Per meglio evidenziare questa peculiarità, consideriamo adesso lo sviluppo in serie di Fourier del segnale *onda triangolare* $x(t)$ rappresentato in Figura 2.17.

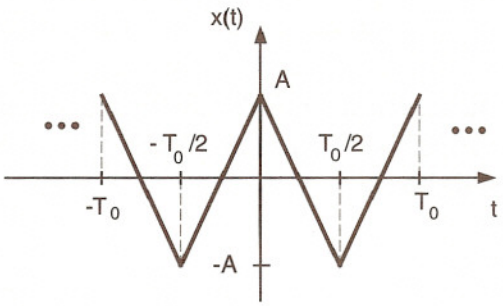


Figura 2.17 Segnale onda triangolare

Questo segnale è pari e alternativo. Il k-esimo coefficiente X_k della sua serie di Fourier può allora essere calcolato sostituendo nella formula semplificata (2.6.9) l'espressione

$$x(t) = A\left(1 - \frac{4t}{T_0}\right) \tag{2.7.4}$$

valida per $0 \le t \le T_0/2$. Si ottiene così

$$\begin{aligned} X_k &= \frac{2}{T_0}\int_0^{T_0/2} A\left(1-\frac{4t}{T_0}\right)\cos(2\pi k f_0 t)\,dt \\ &= \frac{2A}{T_0}\int_0^{T_0/2}\cos(2\pi k f_0 t)\,dt - \frac{8A}{T_0}\int_0^{T_0/2}\frac{t}{T_0}\cos(2\pi k f_0 t)\,dt \end{aligned} \tag{2.7.5}$$

Il calcolo dei due integrali a secondo membro è immediato:

$$\int_0^{T_0/2}\cos(2\pi k f_0 t)\,dt = 0 \tag{2.7.6}$$

$$\int_0^{T_0/2}\frac{t}{T_0}\cos(2\pi k f_0 t)\,dt = \frac{T_0}{(2\pi k)^2}\left[(-1)^k - 1\right] \tag{2.7.7}$$

per cui

$$X_k = \frac{2A}{(\pi k)^2}\left[1-(-1)^k\right] = \begin{cases} 0 & k \text{ pari} \\ \dfrac{4A}{(\pi k)^2} & k \text{ dispari} \end{cases} \tag{2.7.8}$$

Lo spettro del segnale è rappresentato in Figura 2.18. Si nota innanzitutto che anche in questo caso tutti i coefficienti con indice pari sono nulli perché $x(t)$ è alternativo. La caratteristica però più evidente in un confronto con la Figura 2.16 è che l'ampiezza delle armoniche *tende a zero molto più velocemente* quando $k \to \infty$. La (2.7.8) suggerisce infatti che tale velocità è proporzionale a $1/k^2$; quindi, le armoniche superiori hanno meno "importanza" nella sintesi dell'onda triangolare rispetto al caso dell'onda quadra.

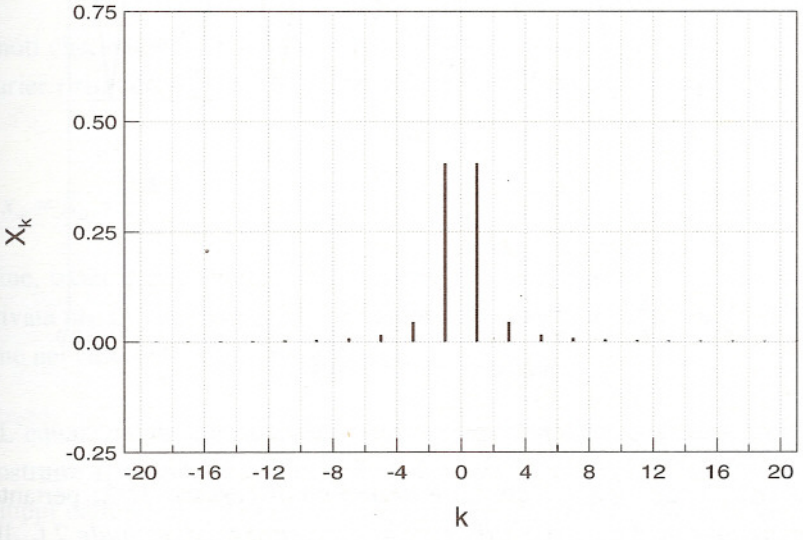


Figura 2.18 Spettro del segnale in Figura 2.17 con $A = 1$

Questo comportamento è un riflesso dell'andamento *temporale* dei due segnali: l'onda quadra presenta *discontinuità*, cioè brusche variazioni temporali del valore del segnale, che viceversa non sono presenti nell'onda triangolare. Possiamo quindi dire che le variazioni "brusche" comportano la presenza di armoniche di ordine più elevato. In altri termini, un segnale avente velocità di cambiamento molto alta necessita per essere ricostruito di molte componenti ad "alta frequenza", cioè di molte armoniche di ordine superiore. Viceversa, un segnale a variazione "più lenta" ha un contenuto di armoniche a frequenze più basse. Questo concetto verrà precisato nel Capitolo 4 discutendo i concetti di *banda* e *durata* di un segnale.

L'andamento delle ampiezze delle armoniche di tipo $1/k^2$ è tipico dei segnali con derivata prima discontinua, come l'onda triangolare di Figura 2.17. A questo proposito, consideriamo un ulteriore esempio.

Esempio 2.5

Calcoliamo i coefficienti dello sviluppo in serie di Fourier del *segnale cosinusoidale raddrizzato a doppia semionda* $x(t) = A\left|\cos(2\pi f_0 t)\right|$ rappresentato in Figura 2.19 ($T_0 = 1/f_0$, $A > 0$).

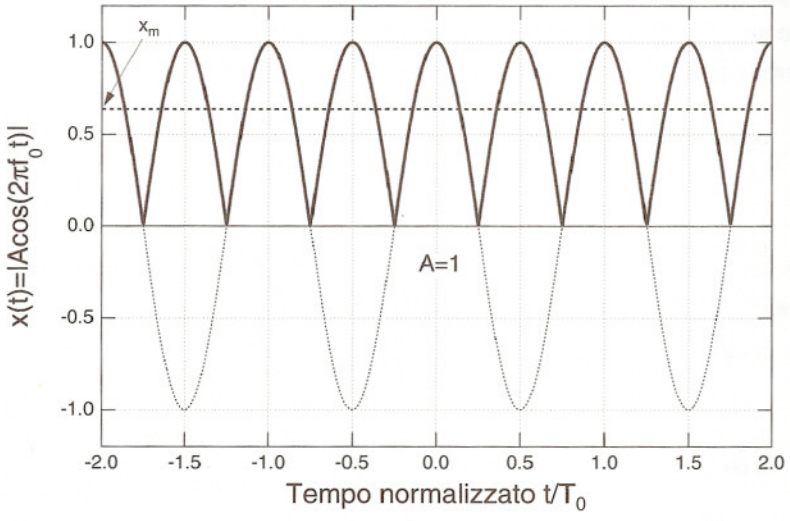


Figura 2.19 Segnale cosinusoidale raddrizzato a doppia semionda

Questo segnale è pari, ma soprattutto è *periodico di periodo* $T_0/2$; pertanto le sue frequenze armoniche sono *multiple della frequenza fondamentale* $2f_0$. Il k-esimo coefficiente X_k della sua serie di Fourier può allora essere calcolato sostituendo l'espressione analitica di $x(t)$ nella formula semplificata (2.6.9) (tenendo conto del fatto che ora il periodo del segnale vale $T_0/2$).

Dunque:

$$X_k = \frac{4A}{T_0}\int_0^{T_0/4} \cos(2\pi f_0 t)\cos(4\pi k f_0 t)\,dt$$

$$= \frac{2A}{T_0}\int_0^{T_0/4} \left\{\cos\left[2\pi(1+2k)f_0 t\right] + \cos\left[2\pi(1-2k)f_0 t\right]\right\} dt \qquad \text{(E2.5.1)}$$

Sviluppando il calcolo dei due integrali si ricava

$$X_k = \frac{A}{\pi}\left\{\frac{1}{1+2k}\sin\left[\frac{\pi}{2}(1+2k)\right] + \frac{1}{1-2k}\sin\left[\frac{\pi}{2}(1-2k)\right]\right\} \qquad \text{(E2.5.2)}$$

Se inoltre si osserva che

$$\sin\left[\frac{\pi}{2}(1+2k)\right]=\sin\left[\frac{\pi}{2}(1-2k)\right]=(-1)^k \tag{E2.5.3}$$

l'espressione di X_k si semplifica come segue:

$$X_k=\frac{A(-1)^k}{\pi}\left\{\frac{1}{1+2k}+\frac{1}{1-2k}\right\}=(-1)^k\frac{2A}{\pi}\frac{1}{1-4k^2} \tag{E2.5.4}$$

Si noti che, essendo $x(t)$ reale e pari, i coefficienti del suo sviluppo in serie di Fourier risultano reali e pari. Inoltre il valor medio del segnale $x(t)$ è espresso da

$$x_m=X_0=\frac{2A}{\pi} \tag{E2.5.5}$$

Infine, osserviamo che si ritrova anche per questo segnale (continuo, ma con derivata prima discontinua) l'andamento delle armoniche proporzionale a $1/k^2$ come nel caso dell'onda triangolare. ❑

L'equazione di sintesi (2.2.9) richiede un numero illimitato di armoniche per ricostruire il segnale periodico $x(t)$. Le considerazioni fatte a proposito degli sviluppi delle onde quadra e triangolare ci suggeriscono però che una approssimazione soddisfacente del segnale può essere conseguita anche con un numero *finito* di armoniche. Per avvalorare questa considerazione ci limiteremo qui a un esempio particolare, consideriando di nuovo il treno di impulsi rettangolari dell'Esempio 2.3. Il suo k-esimo coefficiente di Fourier è

$$X_k=a\,\delta\,\mathrm{sinc}(k\,\delta) \tag{2.7.9}$$

dove, ricordiamo, $\delta=T/T_0$, a è l'ampiezza del segnale, T_0 il suo periodo e T è la durata di ciascun impulso. Poniamo per semplicità $a=1$, $\delta=0.5$ e ricaviamo l'espressione in forma polare della serie di Fourier:

$$x(t)=\frac{1}{2}+\sum_{k=1}^{+\infty}\mathrm{sinc}\left(\frac{k}{2}\right)\cos(2\pi k f_0 t) \tag{2.7.10}$$

D'altronde

$$\operatorname{sinc}\left(\frac{k}{2}\right)=\frac{\sin(k\pi/2)}{k\pi/2}=\begin{cases}0 & k \text{ pari}\\ \dfrac{2}{k\pi}(-1)^{(k-1)/2} & k \text{ dispari}\end{cases} \tag{2.7.11}$$

e quindi la serie può essere semplificata come segue:

$$x(t)=\frac{1}{2}+\frac{2}{\pi}\sum_{\substack{k=1\\ k\ dispari}}^{+\infty}\frac{1}{k}(-1)^{(k-1)/2}\cos(2\pi k f_0 t) \tag{2.7.12}$$

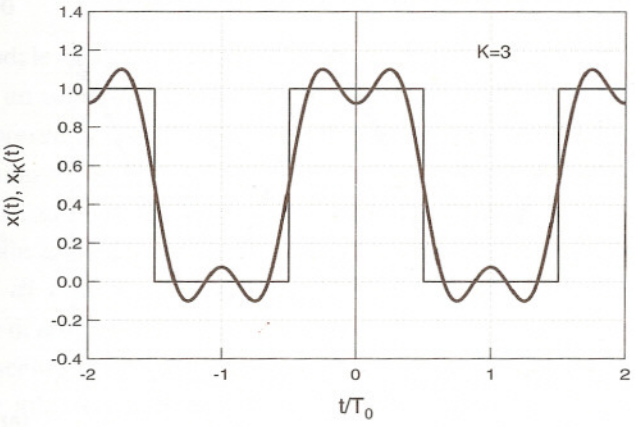
Consideriamo ora il segnale

$$x_K(t)\overset{\Delta}{=}\frac{1}{2}+\frac{2}{\pi}\sum_{\substack{k=1\\ k\ dispari}}^{K}\frac{1}{k}(-1)^{(k-1)/2}\cos(2\pi k f_0 t) \tag{2.7.13}$$

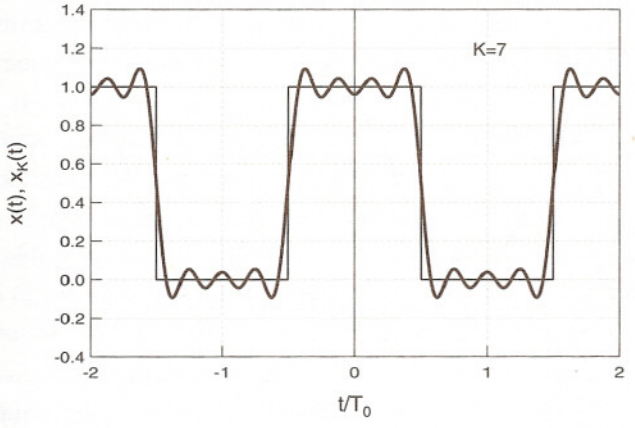
ottenuto *arrestando lo sviluppo* (2.7.12) *alla K-esima armonica.* Nella Figura 2.20 è possibile confrontare il segnale originale (treno di impulsi) $x(t)$ con il segnale approssimante $x_K(t)$ che si ottiene per tre diversi valori del parametro K. Si nota che l'approssimazione del segnale periodico in esame può considerarsi soddisfacente anche con un numero esiguo di armoniche. In corrispondenza dei punti di discontinuità del treno di impulsi, il segnale approssimante presenta inoltre delle fluttuazioni (*ripple*) attorno all'andamento del segnale $x(t)$.

Per il segnale dato, *indipendentemente dal numero di armoniche K* che si usano nella ricostruzione di $x(t)$, si ottiene comunque un segnale approssimante che ha esattamente un valore massimo (nei pressi della discontinuità) pari circa a $1.09\,a$. Questo fatto è noto come *fenomeno di Gibbs*, e la sua presenza fa intuire che la successione di funzioni $\{x_K(t)\}$ non converge *uniformemente* al segnale $x(t)$.

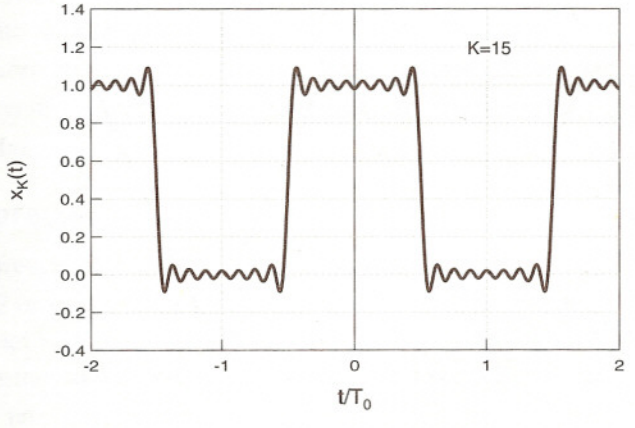
Quanto detto a proposito della rapidità di convergenza dei coefficienti di Fourier dell'onda triangolare di Figura 2.17 è facilmente riscontrabile dal confronto della Figura 2.20 con la Figura 2.21. Quest'ultima mostra ancora il segnale $x_K(t)$ ottenuto per sintesi di sole K armoniche, stavolta però relativamente all'onda triangolare. È evidente che la ricostruzione del segnale periodico originario è molto migliore in questo secondo caso piuttosto che nel caso dell'onda quadra, ovviamente a parità del numero di armoniche considerate.



(a)

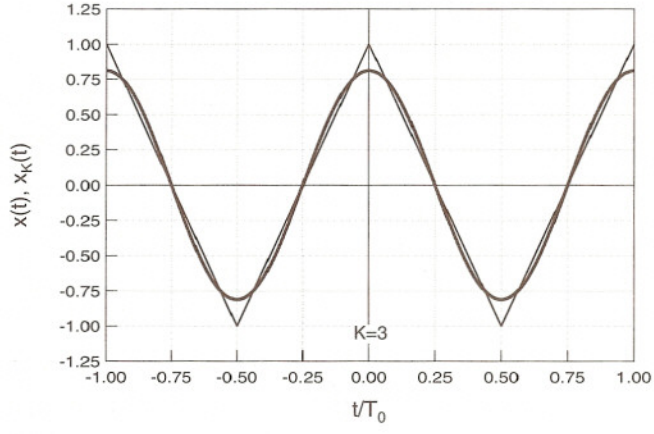


(b)

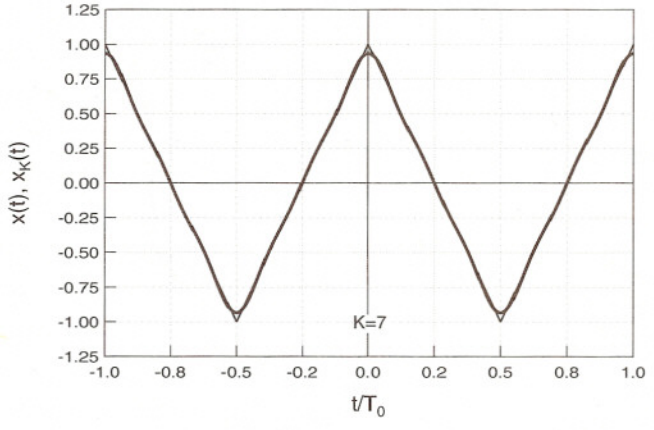


(c)

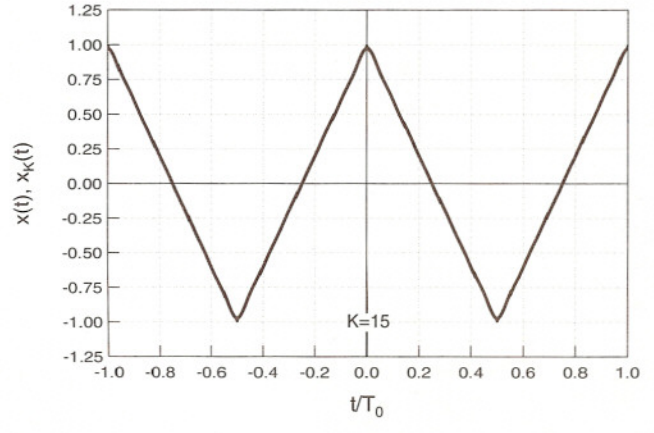
Figura 2.20 Approssimazione del treno di impulsi con 3 (a), 7 (b), e 15 (c) armoniche



(a)



(b)



(c)

Figura 2.21 Approssimazione dell'onda triangolare con 3 (a), 7 (b), e 15 (c) armoniche

Sommario

Il punto nodale di questo capitolo è l'introduzione dello *sviluppo in serie di Fourier* di un segnale a tempo continuo $x(t)$ periodico di periodo T_0. Questo tipo di rappresentazione (l'equazione di *sintesi*) permette di pensare il segnale come "scomposto" in una sovrapposizione di oscillazioni sinusoidali (le *armoniche*) a frequenza multipla della frequenza fondamentale $f_0 = 1/T_0$. L'ampiezza e la fase delle armoniche sono regolate dall'ampiezza e dalla fase dei rispettivi *coefficienti di Fourier* X_k calcolabili a partire dal segnale $x(t)$ attraverso l'equazione di *analisi*.

La conoscenza della successione completa dei coefficienti di Fourier è di fatto equivalente alla conoscenza dell'andamento del segnale nel tempo, cosicché viene usata la scrittura $x(t) \Leftrightarrow X_k$ per rappresentare sinteticamente questa corrispondenza. I coefficienti di Fourier del segnale, in generale complessi, vengono usualmente rappresentati in *modulo* e *fase* in funzione della frequenza armonica cui si riferiscono, e queste rappresentazioni costituiscono i cosiddetti *spettri* di ampiezza e fase del segnale dato. Poiché il segnale $x(t)$ assume valori reali, lo spettro di ampiezza è *simmetrico* rispetto alla frequenza 0 (cioè $|X_k| = |X_{-k}|$) mentre lo spettro di fase è *antisimmetrico* (e cioè $\angle X_k = -\angle X_{-k}$). Altre eventuali proprietà di simmetria del segnale producono ulteriori vincoli sugli spettri di ampiezza e fase del segnale: a segnale pari corrispondono coefficienti di Fourier *reali*, a segnale dispari corrispondono coefficienti *immaginari puri*, a segnale alternativo corrispondono coefficienti di *ordine pari nulli*.

In pratica, la *sintesi* di un segnale periodico viene effettuata con un numero di armoniche *finito*. Questo comporta un certo errore nella ricostruzione del segnale, tanto minore quanto maggiore è il numero di armoniche considerate, ma tanto maggiore quanto più la *velocità di variazione* del segnale è grande: il caso più sfavorevole è quello di un segnale con discontinuità di prima specie, come l'onda quadra.

Esercizi proposti

2.1 Si riprenda in considerazione l'esempio della squadra *R-C* del Paragrafo 2.1. È in grado il lettore di calcolare i coefficienti di Fourier del segnale $y(t)$ nel caso in cui il segnale $x(t)$ sia l'onda quadra di Figura 2.15?

2.2 Dimostrare attraverso la (2.2.9) che se $x(t) \Leftrightarrow X_k$, allora

$$\frac{dx(t)}{dt} \Leftrightarrow j2\pi k f_0 \cdot X_k$$

e sulla base di questo risultato ricavare i coefficienti di Fourier dell'onda quadra di Figura 2.15 da quelli dell'onda triangolare di Figura 2.17.

2.3 Dimostrare attraverso la (2.2.15) che se $x(t) \Leftrightarrow X_k$, allora

$$x(t-t_0) \Leftrightarrow X_k \cdot \exp\{-j2\pi k f_0 t_0\}$$

2.4 Attraverso il risultato (2.7.8) sullo sviluppo in serie dell'onda triangolare, determinare i coefficienti di Fourier del segnale periodico di Figura 2.23.

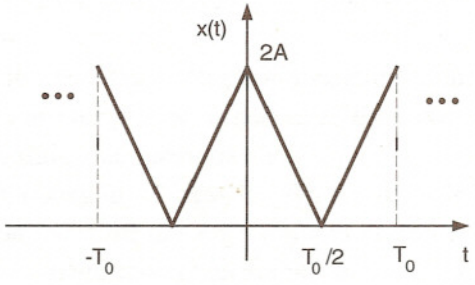


Figura 2.23

2.5 Sfruttando i risultati degli Esercizi 2.2 e 2.4 ricavare i coefficienti di Fourier del segnale di Figura 2.22. [*Suggerimento: si scomponga il segnale dato come differenza tra l'onda triangolare dell'Esercizio 2.4 e ...*]

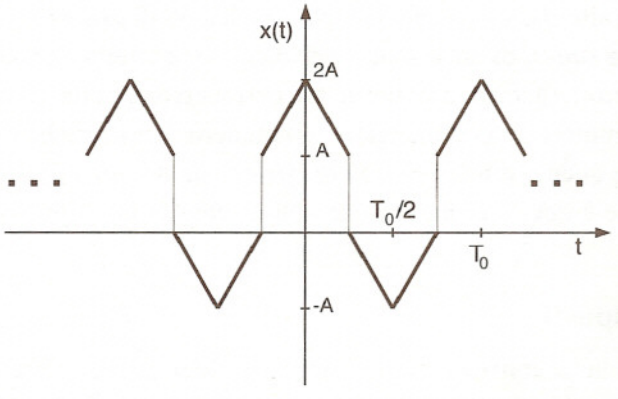
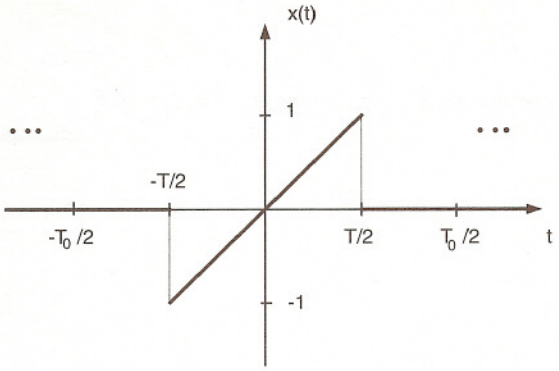
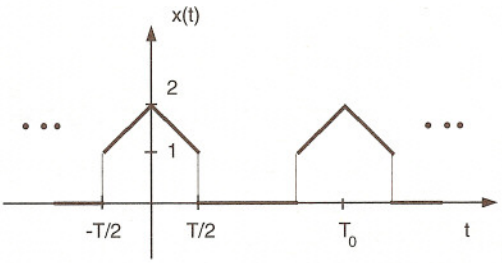


Figura 2.22

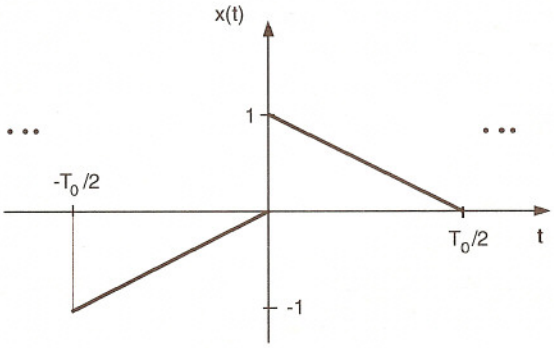
2.6 Determinare l'espressione dei coefficienti delle serie di Fourier dei segnali $x(t)$ periodici di periodo T_0 rappresentati rispettivamente in Figura 2.24*a-b-c* ($T = T_0/2$).



(a)



(b)



(c)

Figura 2.24

CENTRO BIBLIOTECARIO INGEGNERIA PISA

3

Segnali aperiodici a tempo continuo

3.1 Dalla serie all'integrale di Fourier

Il significato e l'importanza della rappresentazione in serie di Fourier di un segnale periodico a tempo continuo sono stati ampiamente discussi nel Capitolo 2. Molti segnali che si osservano nei fenomeni naturali non sono però periodici. Sorge allora immediata la questione della possibilità di ottenere una scomposizione simile alla serie di Fourier *anche per i segnali aperiodici*. È possibile cioè rappresentare anche un segnale non periodico come una opportuna *sovrapposizione* di segnali elementari, in particolare sinusoidali? Per rispondere a questa domanda, consideriamo come caso di studio il segnale aperiodico

$$x(t) = \text{rect}\left(\frac{t}{T}\right) \tag{3.1.1}$$

rappresentato in Figura 3.1, e per ricollegare il discorso a quanto visto nel Capitolo 2 cerchiamo di mettere in relazione questo segnale con il treno di impulsi rettangolari *periodico*

$$x_p(t) = \sum_{n=-\infty}^{+\infty} x(t - nT_0) = \sum_{n=-\infty}^{+\infty} \text{rect}\left(\frac{t - nT_0}{T}\right) \tag{3.1.2}$$

di cui già conosciamo la rappresentazione in serie di Fourier. Come è chiaro, $x_p(t)$ è ottenuto *periodicizzando* $x(t)$ con periodo di ripetizione T_0, come suggerito dalla Figura 3.2. Il segnale originario $x(t)$ può essere considerato come una sorta di caso-limite di un segnale periodico: partendo da $x_p(t)$, si riottiene

l'impulso "base" $x(t)$ centrato in t=0 se si pensa di *fare una periodicizzazione di periodo* $T_0 \to \infty$. Al di là del particolare esempio, se si costruisce un segnale periodico $x_p(t)$ per periodicizzazione del segnale aperiodico $x(t)$ come nella (3.1.2), è vero in generale che

$$x(t) = \lim_{T_0 \to \infty} x_P(t) \tag{3.1.3}$$

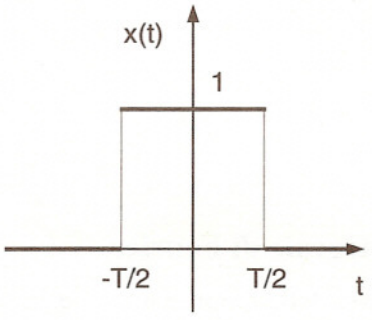


Figura 3.1 Impulso rettangolare aperiodico

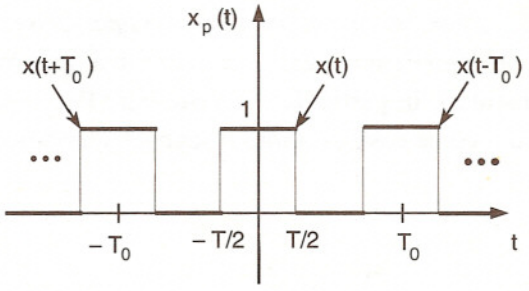


Figura 3.2 Treno periodico di impulsi rettangolari

Naturalmente, il segnale $x_p(t)$, essendo periodico di periodo T_0, può essere rappresentato mediante serie di Fourier:

$$x_p(t) = \sum_{k=-\infty}^{\infty} X_k \, e^{j2\pi k f_0 t} \tag{3.1.4}$$

con $f_0 = 1/T_0$ e con i coefficienti di Fourier X_k dati da

$$X_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x_p(t)\, e^{-j2\pi k f_0 t} dt \qquad (3.1.5)$$

Nasce adesso l'esigenza di stabilire il comportamento della serie di Fourier (3.1.4) e dei relativi coefficienti X_k (3.1.5) quando $T_0 \to \infty$.

Come prima osservazione, è chiaro che aumentando il periodo T_0 di ripetizione si riduce la frequenza fondamentale f_0, e quindi si riduce la differenza tra due generiche frequenze armoniche consecutive $kf_0 - (k-1)f_0 = f_0$. Ciò determina un *infittimento* dello spettro del segnale se la scala di rappresentazione delle frequenze resta la stessa. Inoltre, dalla (3.1.5) si nota che l'*ampiezza* dei coefficienti tende a ridursi man mano che T_0 cresce; al limite, per $T_0 \to \infty$, lo spettro di $x_p(t)$ tende a divenire sempre più fitto e ad assumere valori sempre più piccoli per tutte le frequenze armoniche. La Figura 3.3, relativa al treno di impulsi rettangolari (3.1.2), rappresenta lo spettro di ampiezza del segnale $x_p(t)$ per tre diversi valori del periodo T_0 e per un valore di T assegnato, ed evidenzia chiaramente i due fenomeni appena discussi. Per evitare di dover specificare il particolare valore del periodo T in secondi, la scala delle frequenze è stata *normalizzata* al valore della durata T dell'impulso $x(t)$. Nel grafico si utilizza una variabile "adimensionale" data appunto dal prodotto fT, come chiaramente indicato nella dicitura dell'asse delle ascisse. Questo procedimento di *normalizzazione* delle frequenze o dei tempi (ed eventualmente delle ampiezze) verrà usato sistematicamente nelle pagine a seguire per comodità di rappresentazione.

Si può facilmente ovviare al problema della riduzione delle ampiezze delle righe spettrali definendo, per ciascuna delle frequenze armoniche kf_0, una sorta di "coefficiente di Fourier modificato":

$$X(kf_0) \triangleq T_0\, X_k = \int_{-T_0/2}^{T_0/2} x_p(t)\, e^{-j2\pi k f_0 t} dt = T \operatorname{sinc}(kf_0 T) \qquad (3.1.6)$$

che evidentemente è una quantità che non tende a zero per $T_0 \to \infty$. La Figura 3.4 rappresenta appunto l'andamento del modulo di $X(kf_0)$ nei casi considerati in Figura 3.3. Intenzionalmente, è stata abbandonata la rappresentazione "a righe" dello spettro di ampiezze per passare a una rappresentazione "per punti" che evidenzia solo il valore dello spettro in corrispondenza della frequenza armonica generica, e mette ancora meglio in evidenza l'infittimento delle armoniche.

Riscriviamo dunque l'espansione in serie di Fourier di $x_p(t)$ usando il coefficiente modificato (3.1.6):

$$x_p(t) = \sum_{k=-\infty}^{\infty} X(kf_0)\, e^{j2\pi k f_0 t} \cdot f_0 \tag{3.1.7}$$

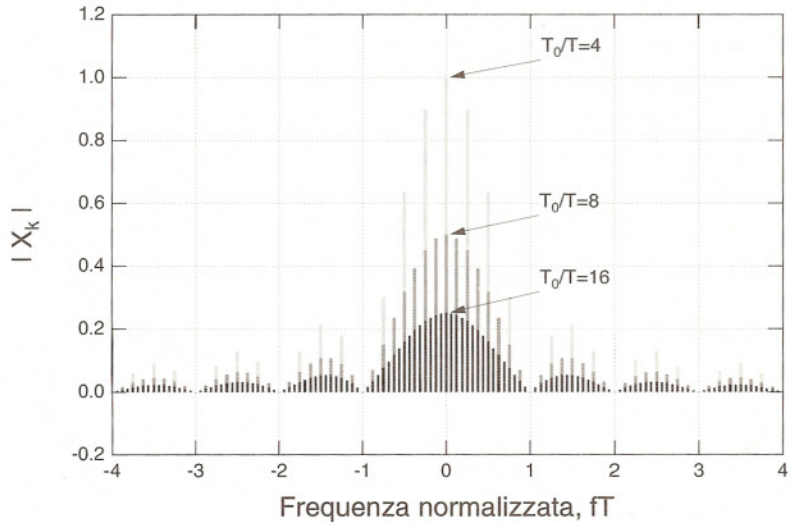


Figura 3.3 Spettro di ampiezza del segnale periodico $x_p(t)$

Possiamo adesso effettuare il passaggio cruciale al limite per $T_0 \to \infty$ (ovvero per $f_0 \to 0$). Il segnale periodico $x_p(t)$ a primo membro della (3.1.7) si trasforma nel segnale aperiodico $x(t)$; si nota inoltre che la serie a secondo membro è esattamente una somma di valori di una funzione valutata sui punti discreti equispaziati kf_0 (e cioè $X(kf_0)\, e^{j2\pi k f_0 t}$), moltiplicati per il valore della distanza f_0 tra due punti consecutivi, *distanza tendente a zero* quando $T_0 \to \infty$. Al limite, la somma (per definizione!) si trasforma in un *integrale*, e si ottiene uno *sviluppo* del segnale aperiodico $x(t)$ come segue:

$$x(t) = \int_{-\infty}^{\infty} X(f)\, e^{j2\pi f t} df \tag{3.1.8}$$

Il segnale aperiodico è dunque rappresentabile attraverso il cosiddetto *integrale di Fourier*. Resta da determinare l'espressione della funzione $X(f)$ che compare nell'integrando della (3.1.8). Innanzitutto, è chiaro che tale quantità risulta una funzione complessa della variabile *continua* f, che mantiene il significato di *frequenza*. L'espressione di $X(f)$ si ottiene passando al limite per $T_0 \to \infty$ nella espressione (3.1.6) del coefficiente di Fourier modificato:

$$X(f) = \lim_{\substack{T_0 \to \infty \\ f_0 \to 0}} \int_{-T_0/2}^{T_0/2} x_p(t)\, e^{-j2\pi k f_0 t}\, dt = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi f t}\, dt \qquad (3.1.9)$$

che rappresenta la *trasformata continua di Fourier* del segnale $x(t)$. In maniera euristica, possiamo dire che la variabile continua f è, in un certo senso, il limite della variabile discreta kf_0 di partenza, quando $f_0 \to 0$. La Figura 3.5, che rappresenta l'ampiezza della $X(f)$ risultante per il treno di impulsi rettangolari, permette di visualizzare il passaggio al limite tra il *coefficiente* di Fourier modificato $X(kf_0)$ di Figura 3.4, ancora funzione di variabile discreta, e la *trasformata* di Fourier $X(f)$, funzione, al contrario, di una variabile continua.

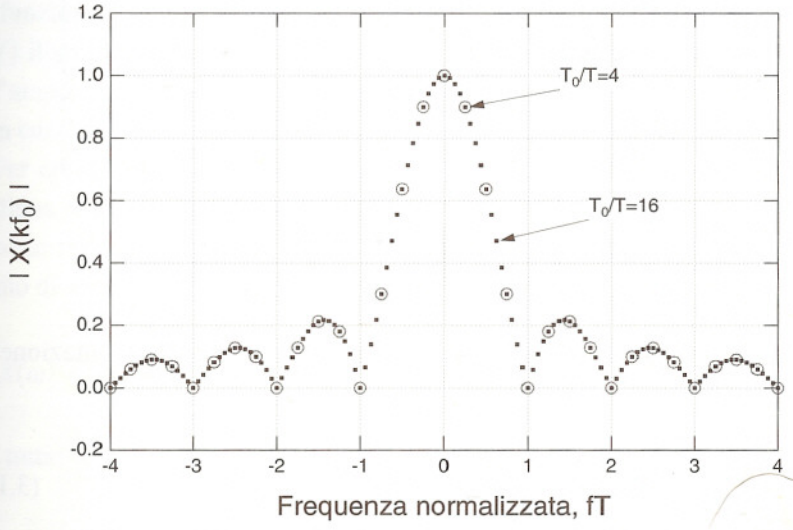


Figura 3.4 Andamento del modulo del coefficiente di Fourier modificato $X(kf_0)$

Commentiamo il risultato ottenuto. Nella serie di Fourier per un segnale periodico, quest'ultimo viene rappresentato mediante componenti sinusoidali a frequenze in *relazione armonica*, cioè tutte multiple di un'unica fondamentale, e di ampiezza *finita*. Nel caso del segnale aperiodico, la (3.1.8), detta anche *antitrasformata di Fourier* (o *trasformata inversa* di Fourier), permette ancora di rappresentare il segnale aperiodico $x(t)$ come la sovrapposizione di componenti sinusoidali, ma stavolta di ampiezza *infinitesima* $X(f)\,df$ e di frequenza f variabile *con continuità* su tutto l'asse reale. In altre parole, il segnale aperiodico è visto come un segnale periodico "di periodo illimitato" e quindi con frequenza fondamentale "infinitamente piccola". La corte discreta di armoniche della serie

degenera quindi nell'insieme *continuo* di componenti proprio dell'integrale di Fourier (antitrasformata).

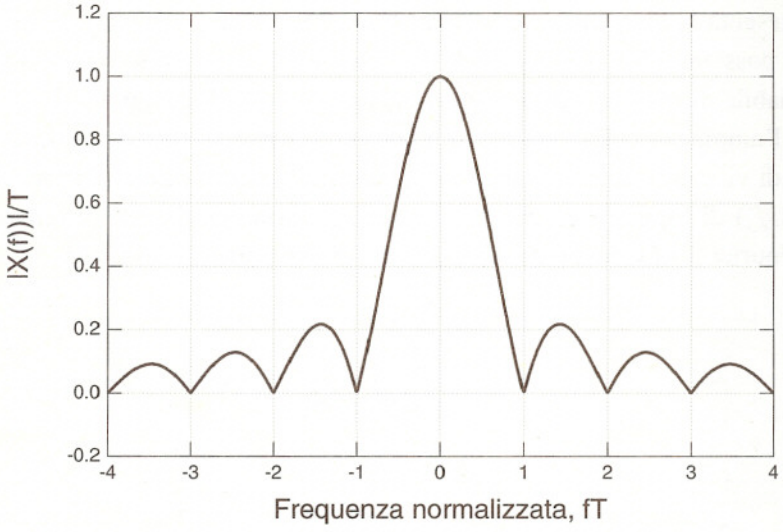


Figura 3.5 Ampiezza della Trasformata di Fourier dell'impulso rettangolare

Riportiamo di nuovo le due equazioni relative alla rappresentazione del segnale aperiodico:

$$x(t) = \int_{-\infty}^{\infty} X(f)\, e^{j2\pi ft}\, df \quad , \quad X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft}\, dt \tag{3.1.10}$$

La prima delle due rappresenta evidentemente un'equazione di *sintesi* che permette di rappresentare il segnale come sovrapposizione di segnali elementari, ed è chiaramente analoga alla (2.2.9) per i segnali periodici; la seconda è un'equazione di *analisi* (analoga alla (2.2.15)) che permette di determinare il *peso* che le varie componenti frequenziali (a tutte le possibili frequenze variabili con continuità da $-\infty$ a $+\infty$) hanno nella composizione di $x(t)$. Tali relazioni mettono in corrispondenza un segnale del tempo con la propria trasformata di Fourier, funzione a valori complessi della frequenza. Come d'uso anche con i coefficienti di Fourier, le relazioni (3.1.10) vengono riassunte con la notazione

$$x(t) \Leftrightarrow X(f) \tag{3.1.11}$$

Un modo alternativo di indicare sinteticamente le operazioni di trasformata e antitrasformata è quello mutuato dalla notazione degli operatori caratteristica dell'analisi funzionale:

$$X(f) = \mathcal{F}[x(t)] \quad , \quad x(t) = \mathcal{F}^{-1}[X(f)] \tag{3.1.12}$$

In analogia a quanto visto per i coefficienti di Fourier X_k, si è soliti estrarre dalla funzione complessa $X(f)$ le funzioni reali modulo $A(f)$ e fase $\vartheta(f)$ secondo la relazione

$$X(f) = A(f)\, e^{j\vartheta(f)} \tag{3.1.13}$$

La funzione $A(f)$ rappresenta lo *spettro di ampiezza* del segnale, la funzione $\vartheta(f)$ il suo *spettro di fase*. Naturalmente, i due spettri forniscono informazioni sull'ampiezza e sulla fase delle componenti frequenziali alla generica frequenza f in cui il segnale viene scomposto dall'operazione di trasformata.

Per completezza, osserviamo che è d'uso talvolta esprimere la trasformata continua di Fourier in funzione della *pulsazione* $\omega = 2\pi f$ (misurata in rad/s) anziché in funzione della frequenza (in Hz). Dalle (3.1.8)-(3.1.9), con un cambiamento di variabile, si possono ricavare le equivalenti relazioni

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t} dt \quad , \quad x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t} d\omega \tag{3.1.14}$$

che tuttavia non verranno più prese in considerazione in questo testo.

Esempio 3.1

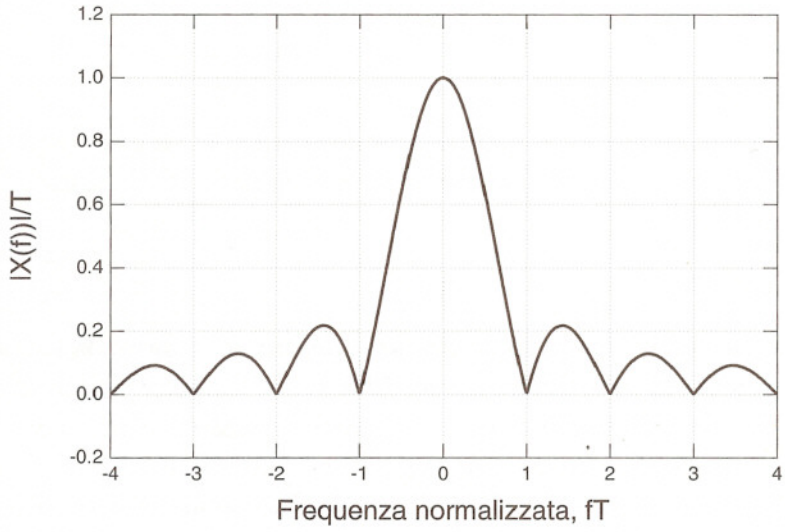
Consideriamo il segnale impulso rettangolare $x(t) = \text{rect}(t/T)$ e calcoliamone la trasformata di Fourier $X(f)$. Questa è data da

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft} dt = \int_{-T/2}^{T/2} e^{-j2\pi ft} dt = \frac{e^{-j2\pi ft}}{-j2\pi f}\bigg|_{-T/2}^{T/2} = \frac{\sin(\pi fT)}{\pi f}$$

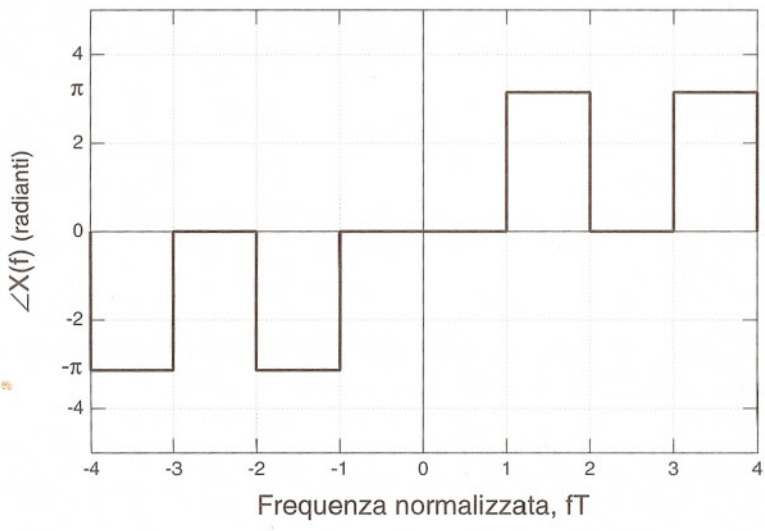
$$= T\,\text{sinc}(fT) \tag{E3.1.1}$$

Gli spettri di ampiezza e di fase del segnale $x(t)$ sono rappresentati nella Figura 3.6. Lo spettro di ampiezza presenta infiniti *nulli* per tutte le frequenze multiple intere dell'*inverso della durata* dell'impulso (esclusa ovviamente $f = 0$). Il risultato trovato si può dunque riassumere come segue:

$$\text{rect}(t/T) \Leftrightarrow T\,\text{sinc}(fT) \qquad \text{(E3.1.2)}$$



(a)



(b)

Figura 3.6 Spettro di ampiezza (a) e di fase (b) dell'impulso rettangolare di ampiezza unitaria

Una verifica sulla correttezza di una coppia *segnale*⇔*trasformata* può essere effettuata osservando che dalla equazione di analisi (3.1.9) con $f = 0$ si ha:

$$X(0)=\int_{-\infty}^{\infty} x(t)\,dt \tag{E3.1.3}$$

cioè il valore dell'integrale su tutto l'asse reale di un segnale è pari al valore per $f=0$ della sua trasformata. Nel nostro caso abbiamo infatti

$$X(0)=T\,\mathrm{sinc}(0)=T \tag{E3.1.4}$$

$$\int_{-\infty}^{\infty}\mathrm{rect}(t/T)\,dt=\int_{-T/2}^{T/2}1\,dt=T \tag{E3.1.5}$$

che confermano la correttezza della (E3.1.1). Naturalmente, la trasformata $X(f)$ (E3.1.1) appena ricavata è esattamente uguale al risultato che si ottiene calcolando il limite del *coefficiente di Fourier modificato* (3.1.6) del treno di impulsi rettangolari per $T_0\to\infty$, cioè $f_0\to 0$, e $kf_0\to f$. ❑

Esempio 3.2

Nella teoria dei segnali e dei sistemi è utile definire la funzione *gradino unitario* $\mathrm{u}(t)$ (detta anche *funzione di Heavyside*), come segue:

$$\mathrm{u}(t)=\begin{cases}1 & t>0\\ 1/2 & t=0\\ 0 & t<0\end{cases} \tag{E3.2.1}$$

Tale funzione, discontinua nell'origine, è utile per rappresentare in modo conciso i segnali cosiddetti *causali* o *cisoidali*, cioè *nulli per* $t<0$. Un esponente di questa classe è il segnale *esponenziale unilatero* $x(t)=e^{-t/T}\,\mathrm{u}(t)$ mostrato in Figura 3.7, la cui trasformata di Fourier $X(f)$ è pari a

$$X(f)=\int_{-\infty}^{\infty}x(t)\,e^{-j2\pi ft}dt=\int_{-\infty}^{\infty}e^{-t/T}\,\mathrm{u}(t)\,e^{-j2\pi ft}dt=\int_{0}^{+\infty}e^{-t(1/T+j2\pi f)}dt$$

$$=-\left.\frac{e^{-t(1/T+j2\pi f)}}{1/T+j2\pi f}\right|_{t=0}^{t=+\infty}=\frac{T}{1+j2\pi fT} \tag{E3.2.2}$$

Allora gli spettri di ampiezza e di fase del segnale $x(t)$ sono dati da

$$A(f)=\frac{T}{\sqrt{1+(2\pi fT)^2}}\quad,\quad \theta(f)=-\arctg(2\pi fT) \tag{E3.2.3}$$

e sono rappresentati in Figura 3.8. Contrariamente alla Figura 3.6, lo spettro di fase è stato rappresentato in gradi anziché radianti, come è d'uso nella pratica.

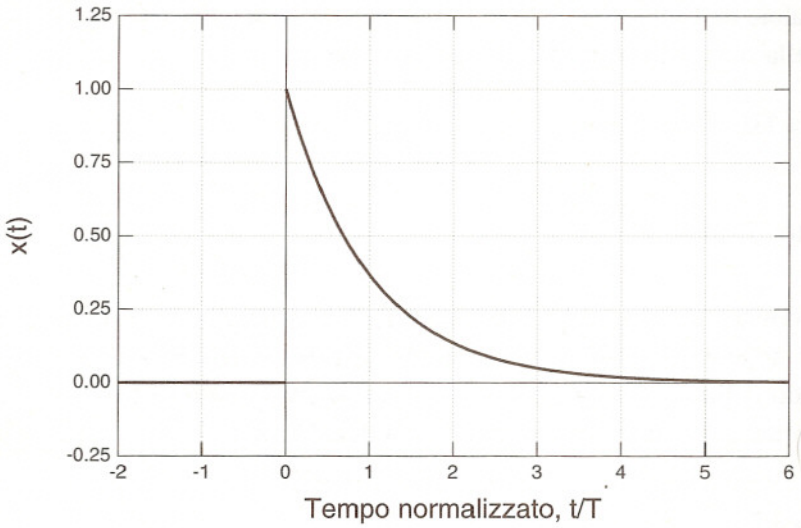


Figura 3.7 Segnale esponenziale unilatero

❑

3.2 Proprietà della trasformata di Fourier

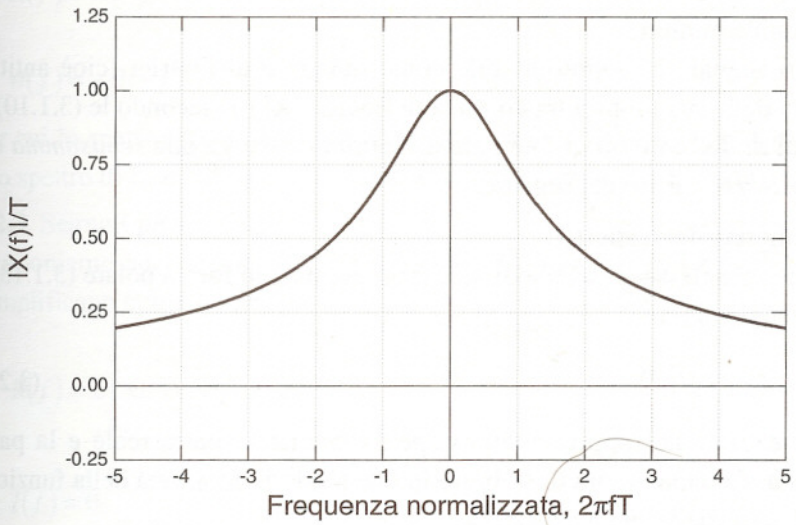
3.2.1 Criteri di esistenza

Indichiamo adesso delle condizioni sufficienti per la rappresentazione del segnale $x(t)$ attraverso la propria trasformata di Fourier $X(f)$, nel senso già discusso riguardo la serie di Fourier. Se tali condizioni sono soddisfatte è possibile affermare che la conoscenza dell'andamento nel tempo del segnale $x(t)$ è equivalente alla conoscenza dell'andamento frequenziale della relativa trasformata di Fourier.

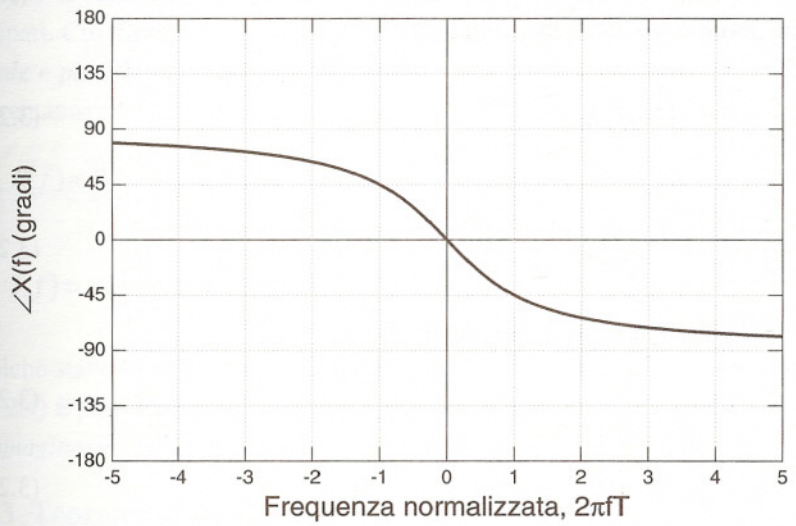
Una prima condizione sufficiente afferma che se il segnale $x(t)$ ha energia finita:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt < +\infty \tag{3.2.1}$$

allora la trasformata $X(f)$ esiste, nel senso che l'integrale nella (3.1.9) è convergente, e la rappresentazione del segnale come integrale di Fourier (antitrasformata) coincide quasi ovunque col segnale originario $x(t)$.



(a)



(b)

Figura 3.8 Spettro di ampiezza (a) e fase (b) del segnale esponenziale unilatero

Un secondo criterio sufficiente meno restrittivo (*criterio di Dirichlet*) può essere enunciato come segue (si confronti con il criterio per la *serie* di Fourier nel Paragrafo 2.3):

- *se* il segnale $x(t)$ è assolutamente sommabile, ovvero $\int_{-\infty}^{\infty} |x(t)|\,dt < +\infty$;
- *se* in qualunque intervallo finito $t_1 \leq t \leq t_2$ il segnale $x(t)$ ha un numero finito di discontinuità di prima specie;

- *se* in qualunque intervallo finito $t_1 \le t \le t_2$ il segnale $x(t)$ ha un numero finito di massimi e minimi;
  allora il segnale è rappresentabile come integrale di Fourier (cioè antitrasformata della sua propria trasformata di Fourier $X(f)$) secondo le (3.1.10), e nei punti di discontinuità l'integrale di Fourier converge alla *semisomma* dei limiti destro e sinistro del segnale.

3.2.2 Simmetrie degli spettri

La funzione complessa $X(f)$ può essere rappresentata in forma polare (3.1.13) o in forma rettangolare:

$$X(f) = R(f) + jI(f) \tag{3.2.2}$$

ove $R(f)$ e $I(f)$ ne rappresentano rispettivamente la parte reale e la parte immaginaria. Cerchiamo ora di stabilire in che modo le proprietà della funzione $x(t)$ si riflettano nella sua trasformata.

Supponiamo che $x(t)$ sia una funzione *reale*; in tal caso le funzioni $R(f)$ e $I(f)$ si ricavano immediatamente dalla (3.1.9):

$$R(f) = \int_{-\infty}^{\infty} x(t)\cos(2\pi ft)dt \tag{3.2.3}$$

$$I(f) = -\int_{-\infty}^{\infty} x(t)\sin(2\pi ft)dt \tag{3.2.4}$$

Da queste espressioni si vede chiaramente che

$$R(f) = R(-f) \tag{3.2.5}$$

$$I(f) = -I(-f) \tag{3.2.6}$$

ovvero la parte reale della trasformata di un segnale reale è una funzione *pari* della frequenza, mentre la parte immaginaria ne è una funzione *dispari*. Queste simmetrie si possono riassumere scrivendo

$$X(f) = X^*(-f) \tag{3.2.7}$$

secondo cui la trasformata di un segnale reale gode della proprietà di *simmetria Hermitiana* (o *coniugata*). La medesima proprietà si riflette ovviamente anche nelle funzioni $A(f)$ e $\vartheta(f)$:

$$A(f) = A(-f) \tag{3.2.8}$$

$$\vartheta(f) = -\vartheta(-f) \tag{3.2.9}$$

per cui lo spettro di ampiezza di un segnale reale è una funzione pari, mentre il suo spettro di fase è dispari (si rivedano a questo proposito le Figure 3.6 e 3.8).

3.2.3 Segnali pari e dispari

Supponiamo adesso che $x(t)$ sia un segnale *reale* e *pari*. Le (3.2.3)-(3.2.4) si semplificano come segue

$$\begin{cases} R(f) = 2\int_0^\infty x(t)\cos(2\pi ft)dt & (3.2.10) \\ I(f) = 0 & (3.2.11) \end{cases}$$

poiché la funzione integranda della (3.2.3) è pari, mentre quella della (3.2.4) è dispari. Ciò dimostra che la trasformata di un segnale reale e pari è una funzione *reale e pari* della frequenza. Se invece $x(t)$ è *reale* e *dispari*, le (3.2.3)-(3.2.4) diventano

$$\begin{cases} R(f) = 0 & (3.2.12) \\ I(f) = -2\int_0^\infty x(t)\sin(2\pi ft)dt & (3.2.13) \end{cases}$$

poiché stavolta la funzione integranda della (3.2.3) è dispari, mentre quella della (3.2.4) è pari. Perciò, la trasformata di un segnale reale e dispari è una funzione *immaginaria pura e dispari*.

3.3 Teoremi sulla trasformata di Fourier

Dalla definizione di trasformata, seguono facilmente alcune ulteriori proprietà, che chiameremo *teoremi*, estremamente utili nel calcolo delle trasformate dei segnali e comunque nell'uso dell'analisi di Fourier in problemi di carattere applicativo.

3.3.1 Teorema di *linearità*

Può essere conveniente in molti casi esprimere un segnale $x(t)$ come combinazione lineare di due segnali $x_1(t)$ e $x_2(t)$:

$$x(t) = a\, x_1(t) + b\, x_2(t) \tag{3.3.1}$$

con a e b costanti. Indicando come di consueto con $X_1(f)$ e $X_2(f)$ le trasformate rispettivamente dei segnali $x_1(t)$ e $x_2(t)$, la trasformata $X(f)$ di $x(t)$ è allora

$$X(f) = a\, X_1(f) + b\, X_2(f) \tag{3.3.2}$$

Infatti, applicando la definizione di trasformata (3.1.9) abbiamo

$$X(f) = \int_{-\infty}^{\infty} x(t)\, e^{-j2\pi ft}\, dt = \int_{-\infty}^{\infty} \left[a\, x_1(t) + b\, x_2(t)\right] e^{-j2\pi ft}\, dt \tag{3.3.3}$$

e sfruttando la linearità dell'integrale otteniamo

$$X(f) = a \int_{-\infty}^{\infty} x_1(t)\, e^{-j2\pi ft}\, dt + b \int_{-\infty}^{\infty} x_2(t)\, e^{-j2\pi ft}\, dt = a\, X_1(f) + b\, X_2(f) \tag{3.3.4}$$

3.3.2 Teorema di *dualità*

La similitudine tra le relazioni (3.1.10) di trasformata e antitrasformata intese come "operatori" sulle funzioni rispettivamente $x(t)$ e $X(f)$ permette di risolvere una questione: se $X(f)$ indica la trasformata del segnale $x(t)$, qual è la trasformata del segnale *temporale* $X(t)$, avente cioè lo stesso andamento temporale originariamente posseduto nell' ambito frequenziale dalla trasformata di $x(t)$? La risposta è la seguente: se

$$x(t) \Leftrightarrow X(f) \tag{3.3.5}$$

allora

$$X(t) \Leftrightarrow x(-f) \tag{3.3.6}$$

Infatti, sappiamo che il segnale $x(t)$ è legato alla sua trasformata dalla relazione

$$x(t) = \int_{-\infty}^{\infty} X(f)\, e^{j2\pi ft}\, df \tag{3.3.7}$$

Scambiando formalmente le variabili t ed f nella (3.3.7) si ricava

$$x(f) = \int_{-\infty}^{\infty} X(t)\, e^{j2\pi ft}\, dt \tag{3.3.8}$$

Se poi in questa relazione si effettua un cambiamento di variabile sostituendo alla variabile f la variabile $-f$, si ottiene

$$x(-f) = \int_{-\infty}^{\infty} X(t)\, e^{-j2\pi ft}\, dt \tag{3.3.9}$$

che coincide con la (3.3.6).

Esempio 3.3

Consideriamo il segnale $x(t) = \text{sinc}(B\,t)$. Dall' Esempio 3.1 sappiamo che

$$\text{rect}\left(\frac{t}{T}\right) \Leftrightarrow T\ \text{sinc}(fT) \tag{E3.3.1}$$

Utilizzando la proprietà di dualità è allora possibile scrivere:

$$T\text{sinc}(T\,t) \Leftrightarrow \text{rect}\left(-\frac{f}{T}\right) \tag{E3.3.2}$$

Poiché l'impulso rettangolare è una funzione pari, ponendo formalmente $B = 1/T$ si trova la seguente relazione, duale della (E3.1.1):

$$\text{sinc}(B\,t) \Leftrightarrow \frac{1}{B}\,\text{rect}\left(\frac{f}{B}\right) \tag{E3.3.3}$$

❑

3.3.3 Teorema del *ritardo*

Come viene modificata la trasformata di un segnale se questo viene *traslato* sull'asse dei tempi (cioè, anticipato o ritardato)? Sia dunque $X(f)$ la trasformata del segnale $x(t)$; allora la trasformata del segnale traslato verso destra della quantità t_0 è

$$x(t - t_0) \Leftrightarrow X(f)\, e^{-j2\pi ft_0} \tag{3.3.10}$$

Questa operazione corrisponde evidentemente a un ritardo se $t_0 > 0$, e a un anticipo se $t_0 < 0$. Applicando la definizione di trasformata si ha

$$x(t - t_0) \Leftrightarrow \int_{-\infty}^{\infty} x(t - t_0)\, e^{-j2\pi ft}\, dt \tag{3.3.11}$$

Se nell'integrale che figura nella relazione precedente si effettua il cambiamento di variabile $\alpha = t - t_0$ si ricava

$$x(t-t_0) \Leftrightarrow \int_{-\infty}^{\infty} x(\alpha)\, e^{-j2\pi f(\alpha+t_0)} d\alpha = e^{-j2\pi f t_0} \int_{-\infty}^{\infty} x(\alpha)\, e^{-j2\pi f \alpha} d\alpha \tag{3.3.12}$$

dalla quale segue immediatamente la (3.3.10).

Questa proprietà mostra che un ritardo temporale modifica lo spettro di fase della trasformata del segnale ma *non cambia il suo spettro di ampiezza.* Infatti, se si indica con $Y(f)$ la trasformata di Fourier del segnale $x(t-t_0)$, il teorema del ritardo si traduce nelle seguenti relazioni:

$$|Y(f)| = |X(f)| \tag{3.3.13}$$

$$\angle Y(f) = \angle X(f) - 2\pi f t_0 \tag{3.3.14}$$

la seconda delle quali sottolinea che lo sfasamento introdotto dal ritardo t_0 varia linearmente con la frequenza.

Esempio 3.4

Consideriamo l'impulso rettangolare dell'Esempio 3.1 ritardato di metà della sua durata (Figura 3.9):

$$x(t) = \text{rect}\left(\frac{t - T/2}{T}\right) \tag{E3.4.1}$$

e calcoliamo la sua trasformata continua.

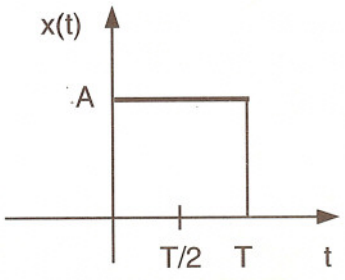


Figura 3.9 Impulso rettangolare ritardato

Utilizzando la (E3.1.1) e il teorema del ritardo (3.3.10) si ricava che

$$X(f) = T\,\mathrm{sinc}(fT)\,e^{-j2\pi fT/2} = AT\,\mathrm{sinc}(fT)\,e^{-j\pi fT} \qquad \text{(E3.4.2)}$$

Lo spettro di ampiezza del segnale è ancora quello rappresentato in Figura 3.5, mentre il nuovo spettro di fase è rappresentato in Figura 3.10. Si noti che lo spettro di fase viene rappresentato riducendo il valore della fase tra $-\pi$ e π, cioè se ne considera il valore *modulo-2π*.

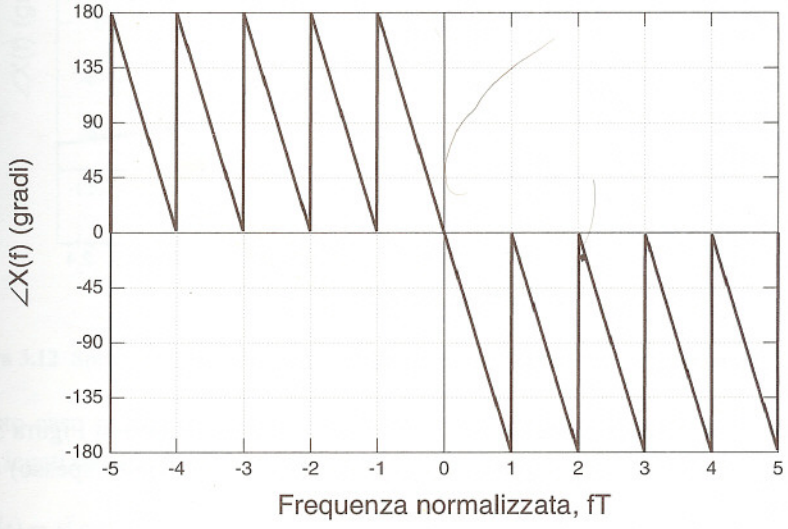


Figura 3.10 Spettro di fase dell'impulso rettangolare ritardato

❑

Esempio 3.5

Consideriamo il segnale esponenziale unilatero dell'Esempio 3.2 ritardato di $t_0 = T/16$ secondi:

$$x(t) = e^{-(t-t_0)/T}\,\mathrm{u}(t-t_0) \qquad \text{(E3.5.1)}$$

e calcoliamone la trasformata continua $X(f)$. Il segnale $x(t)$ è rappresentato a tratto spesso in Figura 3.11, insieme al segnale esponenziale non ritardato, mostrato a linea tratteggiata. Utilizzando la (E3.2.2) e il teorema del ritardo (3.3.10) si ricava:

$$X(f) = \frac{T}{1+j2\pi fT}\,e^{-j2\pi ft_0} = \frac{T}{1+j2\pi fT}\,e^{-j\pi fT/8} \qquad \text{(E3.5.2)}$$

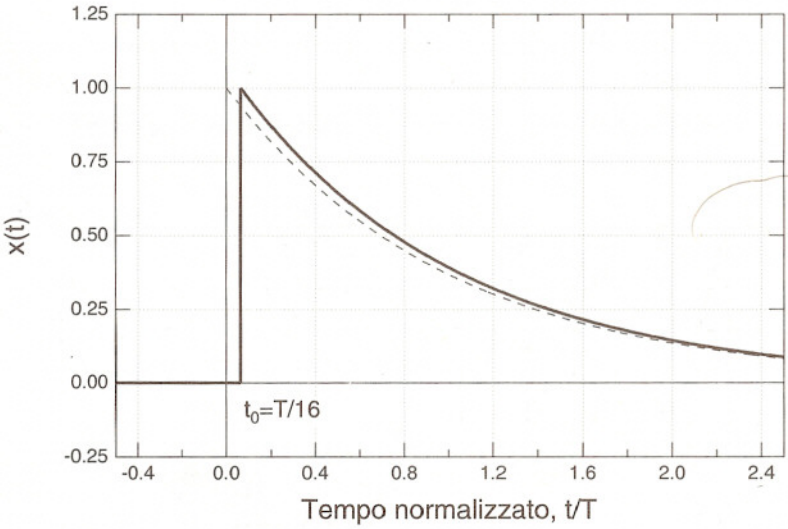


Figura 3.11 Segnale esponenziale ritardato di t_0=T/16

Lo spettro di ampiezza del segnale è ancora quello rappresentato in Figura 3.8a. Lo spettro di fase è invece rappresentato in Figura 3.12 (a tratto spesso) ed è espresso da

$$\theta(f) = -\operatorname{arctg}(2\pi fT) - \pi fT/8 \qquad \text{(E3.5.3)}$$

Nella stessa figura sono anche illustrati gli andamenti dei due addendi a secondo membro della (E3.5.3). ❑

3.3.4 Teorema del *cambiamento di scala*

Per illustrare il significato di questo teorema facciamo riferimento come caso di studio all'impulso rettangolare $x(t)$ di durata T (3.1.1), avente come trasformata $X(f) = T\,\mathrm{sinc}(fT)$, e a un impulso rettangolare $y(t)$ di *durata doppia*:

$$y(t) = \mathrm{rect}\left(\frac{t}{2T}\right) \Leftrightarrow 2T\,\mathrm{sinc}(2fT) \qquad (3.3.15)$$

Possiamo considerare $y(t)$ come una versione "rallentata" di $x(t)$, perché la forma del segnale è rimasta inalterata, ma la sua durata è *raddoppiata*. Dalla (3.3.15) si nota che $y(t)$ ha ancora una trasformata del tipo $\mathrm{sinc}(\cdot)$, il cui lobo principale però ha una larghezza (da nullo a nullo) pari a $1/T$, cioè pari *alla metà* della larghezza $2/T$ del lobo principale della trasformata di $x(t)$.

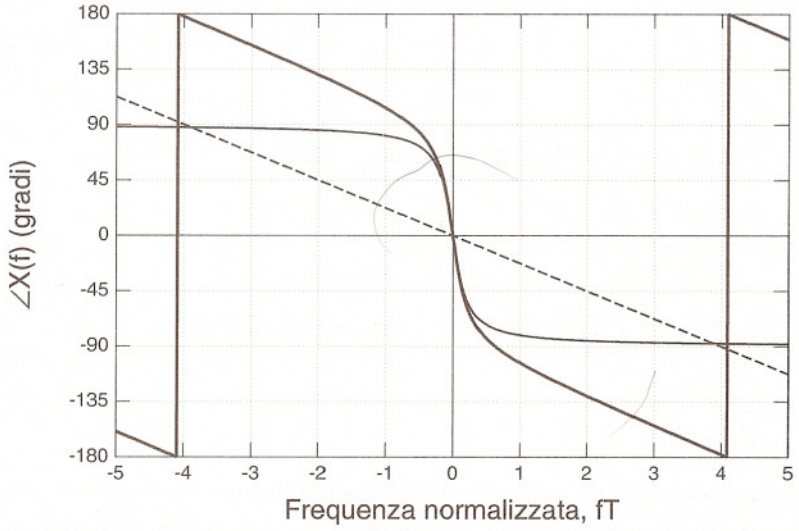


Figura 3.12 Spettro di fase dell'esponenziale unilatero ritardato di $T/16$

Questo caso particolare esemplifica la situazione generale in cui i due segnali sono legati dalla relazione

$$y(t) = x(\alpha t) \tag{3.3.16}$$

cioè si effettua un *cambiamento della scala temporale*. Moltiplicando la variabile indipendente t del segnale $x(t)$ per il coefficiente α si producono i seguenti effetti:

$|\alpha| > 1 \quad \Rightarrow$ *compressione* della scala dei tempi
$|\alpha| < 1 \quad \Rightarrow$ *dilatazione* della scala dei tempi
$\alpha < 0 \quad \Rightarrow$ *inversione* della scala dei tempi

In altri termini, se $|\alpha| < 1$ l'evoluzione del segnale viene "rallentata", viceversa se $|\alpha| > 1$ il segnale viene "accelerato". Operazioni di questo tipo vengono effettuate correntemente nell'elaborazione dei segnali *registrando* il segnale su nastro o su disco a una certa velocità e *riproducendolo* a velocità diversa. Supponiamo per semplicità che sia $\alpha > 0$ e calcoliamo la trasformata del segnale $x(\alpha t)$. Possiamo scrivere:

$$x(\alpha t) \Leftrightarrow \int_{-\infty}^{\infty} x(\alpha t)\, e^{-j2\pi ft}\, dt \tag{3.3.17}$$

Effettuando la sostituzione $z = \alpha\, t$ a secondo membro si ha

$$x(\alpha t) \Leftrightarrow \frac{1}{\alpha}\int_{-\infty}^{\infty} x(z)\, e^{-j2\pi f z/\alpha}\, dz = \frac{1}{\alpha} X\left(\frac{f}{\alpha}\right) \quad , \quad \alpha > 0 \tag{3.3.18}$$

Se invece è $\alpha < 0$ allora la (3.3.18) si modifica nella

$$x(\alpha t) \Leftrightarrow -\frac{1}{\alpha}\int_{-\infty}^{\infty} x(z)\, e^{-j2\pi f z/\alpha}\, dz = -\frac{1}{\alpha} X\left(\frac{f}{\alpha}\right) \quad , \quad \alpha < 0 \tag{3.3.19}$$

I risultati (3.3.18)-(3.3.19) possono allora essere riassunti da

$$x(\alpha t) \Leftrightarrow \frac{1}{|\alpha|} X\left(\frac{f}{\alpha}\right) \tag{3.3.20}$$

Si nota che una *dilatazione* dell'asse dei tempi comporta una *compressione* dell'asse delle frequenze, e viceversa. Se infatti il segnale viene "rallentato" vengono a predominare le componenti frequenziali a bassa frequenza che sono responsabili per così dire dell'evoluzione lenta del segnale; lo spettro allora si "addensa" nell'intorno della frequenza nulla.

Esempio 3.6

Consideriamo i seguenti due segnali (Figura 3.13):

$$x(t) = \mathrm{sinc}\left(\frac{t}{T}\right) \tag{E3.6.1}$$

$$y(t) = \mathrm{sinc}\left(\frac{t}{2T}\right) \tag{E3.6.2}$$

Se si osserva che

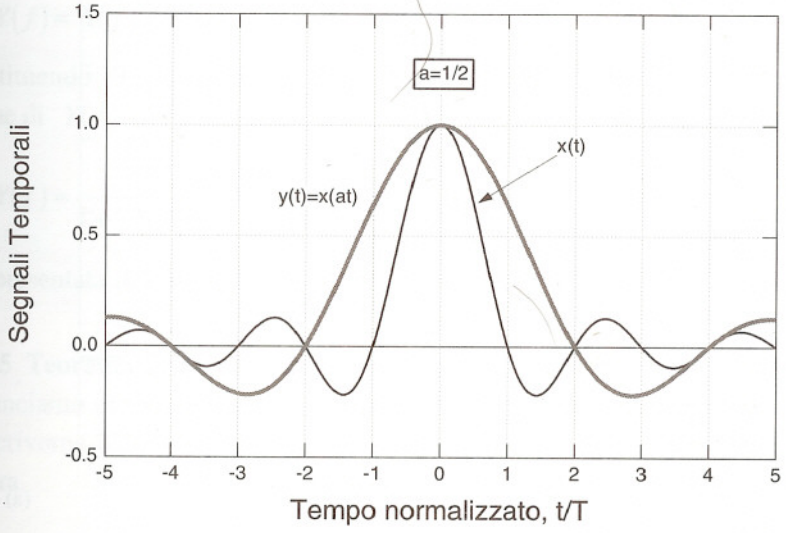
$$y(t) = x(at) \tag{E3.6.3}$$

con $\alpha = 1/2$, applicando il teorema del cambiamento di scala si ricava:

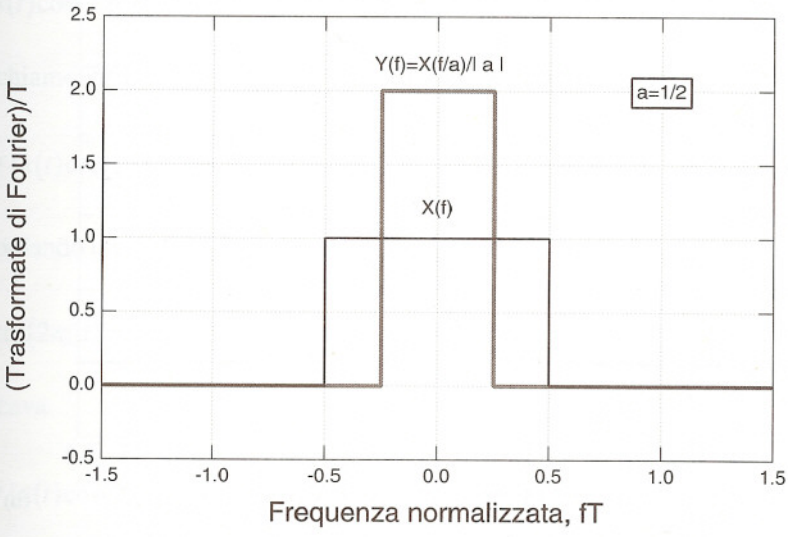
$$Y(f) = 2\, X(2f) = 2T\, \mathrm{rect}(2fT) \tag{E3.6.4}$$

L'andamento dei segnali $x(t)$ e $y(t)$ e delle loro trasformate $X(f)$ e $Y(f)$ è illustrato nella Figura 3.13, che evidenzia molto bene il comportamento "duale" nei domini tempo e frequenza: al segnale più "veloce" $x(t)$ (compresso nel

tempo) corrisponde una trasformata $X(f)$ contenente componenti a più alte frequenze (espansa in frequenza).



(a)



(b)

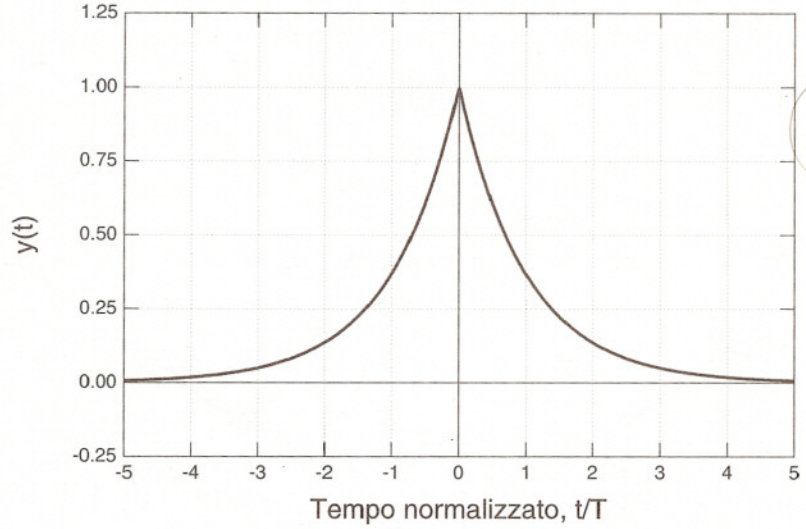
Figura 3.13 Andamento temporale (a) e spettro (b) dei segnali dell'Esempio 3.6

❑

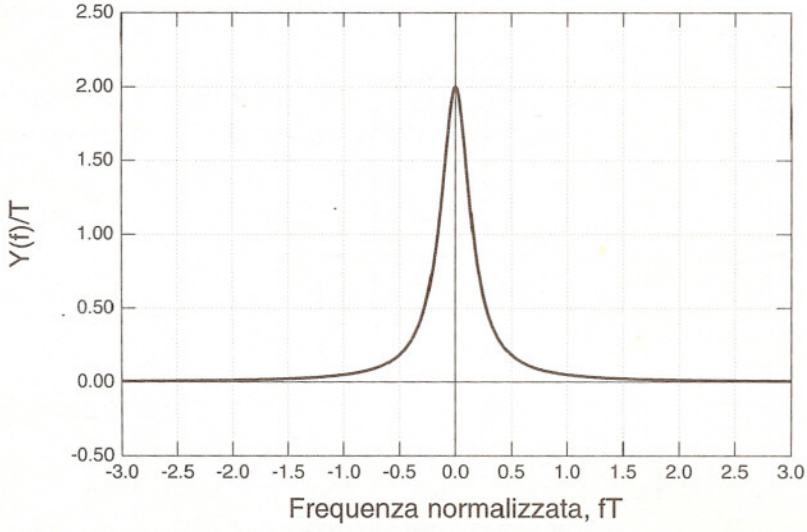
Esempio 3.7

Calcoliamo la trasformata di Fourier $Y(f)$ del segnale *esponenziale bilatero* rappresentato in Figura 3.14a:

$$y(t) = e^{-|t|/T} \tag{E3.7.1}$$



(a)



(b)

Figura 3.14 Andamento temporale (a) e spettro (b) del segnale esponenziale bilatero

Il segnale $y(t)$ può essere espresso in funzione del segnale esponenziale unilatero $x(t) = e^{-t/T}u(t)$ dell'Esempio 3.2 come

$$y(t) = x(t) + x(-t) \tag{E3.7.2}$$

Utilizzando il teorema del cambiamento di scala con $\alpha = -1$, dalla (E3.7.2) si ricava:

$$Y(f) = X(f) + X(-f) \tag{E3.7.3}$$

Sostituendo l'espressione di $X(f)$ data dalla (E3.7.2) si trova infine l'espressione di $Y(f)$:

$$Y(f) = \frac{T}{1 + j2\pi fT} + \frac{T}{1 - j2\pi fT} = \frac{2T}{1 + (2\pi fT)^2} \tag{E3.7.4}$$

rappresentata in Figura 3.14b. ❑

3.3.5 Teorema della *modulazione*

Enunciamo ora formalmente il cosiddetto teorema della modulazione prima di descriverne una applicazione importante. Se, come di consueto, $x(t) \Leftrightarrow X(f)$, allora

$$x(t)\cos(2\pi f_0 t) \Leftrightarrow \frac{X(f - f_0) + X(f + f_0)}{2} \tag{3.3.21}$$

Cerchiamo infatti la trasformata del segnale a primo membro:

$$\mathcal{F}[x(t)\cos(2\pi f_0 t)] = \int_{-\infty}^{\infty} x(t)\cos(2\pi f_0 t)\, e^{-j2\pi ft}\, dt \tag{3.3.22}$$

Ricordando che

$$\cos(2\pi f_0 t) = \frac{e^{j2\pi f_0 t} + e^{-j2\pi f_0 t}}{2} \tag{3.3.23}$$

si ricava

$$\mathcal{F}[x(t)\cos(2\pi f_0 t)] = \frac{1}{2}\int_{-\infty}^{\infty} x(t)\, e^{j2\pi f_0 t} e^{-j2\pi ft}\, dt + \frac{1}{2}\int_{-\infty}^{\infty} x(t)\, e^{-j2\pi f_0 t} e^{-j2\pi ft}\, dt =$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} x(t)\, e^{-j2\pi (f - f_0)t}\, dt + \frac{1}{2}\int_{-\infty}^{\infty} x(t)\, e^{-j2\pi (f + f_0)t}\, dt \tag{3.3.24}$$

Rivolgiamo ora la nostra attenzione al primo integrale della (3.3.24); è immediato rendersi conto che

$$\int_{-\infty}^{\infty} x(t)\, e^{-j2\pi(f-f_0)t}\, dt = X(f-f_0) \tag{3.3.25}$$

Una prima conclusione che possiamo trarre è la seguente: se un segnale viene moltiplicato per un fattore *esponenziale complesso* $e^{j2\pi f_0 t}$, la sua trasformata di Fourier viene traslata attorno alla frequenza f_0. Questo risultato rappresenta la cosiddetta proprietà di *traslazione in frequenza* della trasformata e può essere riassunto come segue:

$$x(t)\, e^{j2\pi f_0 t} \Leftrightarrow X(f-f_0) \tag{3.3.26}$$

Allora la trasformata cercata del segnale *modulato* $x(t)\cos(2\pi f_0 t)$ può essere espressa come

$$\mathcal{F}[x(t)\cos(2\pi f_0 t)] = \frac{X(f-f_0)+X(f+f_0)}{2} \tag{3.3.27}$$

Esaminiamo ora una applicazione pratica di questo risultato.

Esempio 3.8

È noto che gli oggetti metallici *riflettono* le onde elettromagnetiche, e che questa proprietà dei corpi conduttori si accentua al crescere della *frequenza* della radiazione incidente. Su questo fenomeno sono basati i sistemi radar[1] di rivelazione e misura della distanza di oggetti metallici (tipicamente aerei o navi). Lo schema di principio di un radar è rappresentato in Figura 3.15: l'antenna trasmittente emette un *impulso radio* $x(t)$ e si dispone a ricevere poi un segnale $y(t)$ di *eco* riemesso dal bersaglio per semplice riflessione. L'eventuale ricezione di un'eco permette di *rivelare la presenza* del "bersaglio radar", cioè di un oggetto metallico di grosse dimensioni.

Attraverso la misura del *ritardo* dell'eco rispetto all'istante di trasmissione dell'impulso è possibile inoltre calcolare la distanza dell'oggetto riflettente dal trasmettitore. I segnali trasmesso e ricevuto sono schematizzati in Figura 3.16. Se indichiamo con τ il ritardo di ricezione dell'eco radar e d la distanza che separa il radar dall'oggetto individuato, possiamo scrivere

$$d = c\tau/2 \tag{E3.8.1}$$

[1] *Radar* è l'acronimo di *RAdio Detection And Ranging*, letteralmente: rivelazione e misura della distanza tramite onde radio.

ove c è la velocità di propagazione dell'onda elettromagnetica ($c = 3 \cdot 10^8$ m/s). Il fattore 1/2 che compare nella relazione (E3.8.1) deriva dal fatto che l'impulso percorre *due volte* la distanza d antenna-bersaglio.

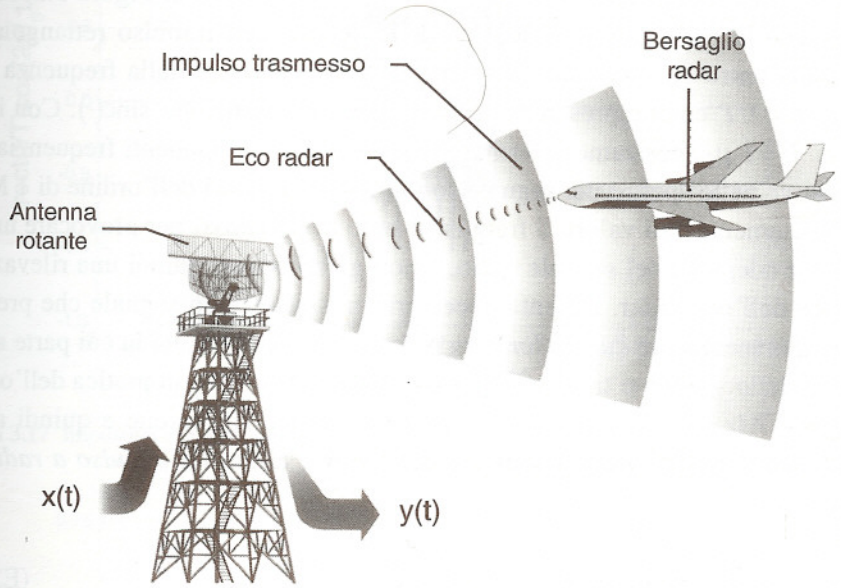


Figura 3.15 Sistema radar

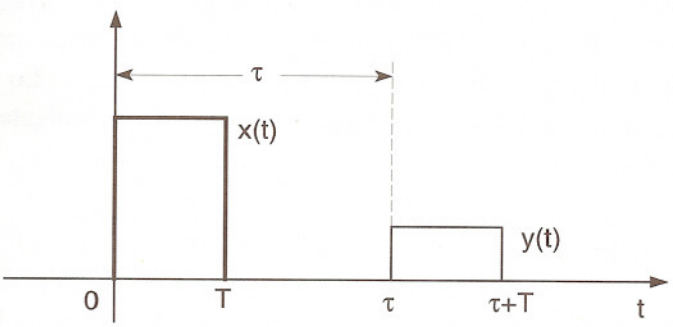


Figura 3.16 Segnale trasmesso ed eco radar

Il valore tipico della durata T dell'impulso trasmesso $x(t)$ è pari all'incirca a 1 μs. Viene utilizzato un valore così piccolo per evitare che, usando impulsi di durata maggiore, l'eco possa *sovrapporsi* all'impulso trasmesso. Lo spettro di

ampiezza del segnale trasmesso è allora

$$|X(f)| = T\,|\mathrm{sinc}(fT)| \tag{E3.8.2}$$

Se si osserva l'andamento della funzione $|X(f)|$ riportato in Figura 3.5, si può notare che la parte più significativa dello spettro dell'impulso rettangolare è contenuta entro un intervallo frequenziale che si estende dalla frequenza zero fino a $f_1 = 1/T$ che rappresenta il primo nullo della funzione $\mathrm{sinc}(\cdot)$. Con il valore di $T = 1$ μs, possiamo affermare dunque che le componenti frequenziali significative nello spettro del segnale sono limitate a valori dell'ordine di 1 MHz. Sfortunatamente, tali valori di frequenza sono troppo *bassi* per provocare una riflessione adeguata del segnale radio, e non garantiscono quindi una rilevazione efficace dell'eco radar. Diventa necessario trasmettere un segnale che presenti contemporaneamente *durata limitata* per quanto già visto, ma la cui parte significativa dello spettro si trovi a *frequenze molto più elevate*, in pratica dell'ordine di $f_0 = 1\,\mathrm{GHz} = 10^9\,\mathrm{Hz}$, per dar luogo a una buona riflessione e quindi a una buona eco. Ciò che viene trasmesso dal radar è allora un *impulso a radiofrequenza*

$$x(t) = \mathrm{rect}\left(\frac{t}{T}\right)\cos(2\pi f_0 t) \tag{E3.8.3}$$

il cui andamento è rappresentato in Figura 3.17. Il segnale risultante è uno "spezzone" di durata T dell'oscillazione cosiddetta *portante* alla frequenza f_0. Nella Figura 3.17, la frequenza portante è pari a sole 10 volte l'inverso della durata dell'impulso. In realtà, se $T = 1$ μs ed $f_0 = 1$ GHz, l'impulso trasmesso risulta formato da un *migliaio* di cicli dell'oscillazione portante. Lo spettro dell'impulso a radiofrequenza è immediatamente calcolabile mediante il *teorema della modulazione:*

$$\mathrm{rect}\left(\frac{t}{T}\right)\cos(2\pi f_0 t) \Leftrightarrow \frac{T\,\mathrm{sinc}\left[(f - f_0)T\right] + T\,\mathrm{sinc}\left[(f + f_0)T\right]}{2} \tag{E3.8.4}$$

che viene rappresentato, ancora nel caso semplificato $f_0 = 10/T$, in Figura 3.18. Si ottiene uno spettro sostanzialmente concentrato attorno alla frequenza f_0, mantenendo un segnale di tipo impulsivo con durata ancora pari a T. Nel caso reale in cui $f_0 = 1$ GHz, le due repliche dello spettro originario create dalla modulazione sono molto più "distanziate" sull'asse frequenziale nei confronti della larghezza dei lobi delle funzioni $\mathrm{sinc}(\cdot)$ di quanto la Figura 3.18 non mostri.

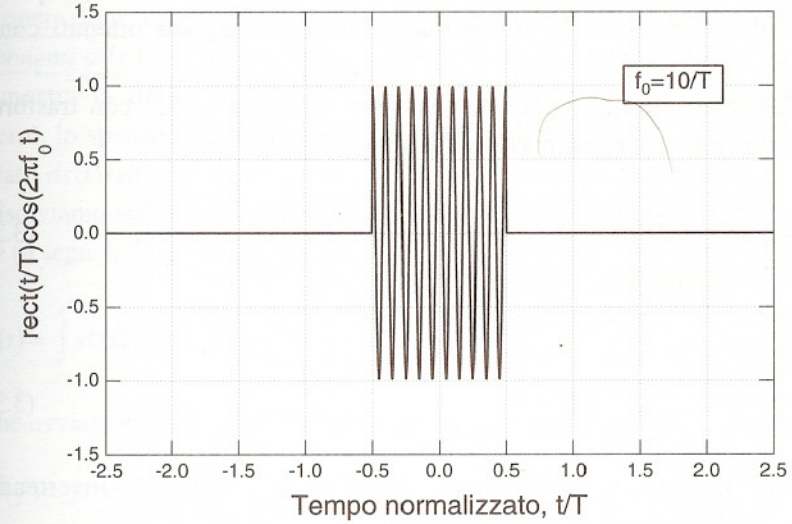


Figura 3.17 Impulso a radiofrequenza

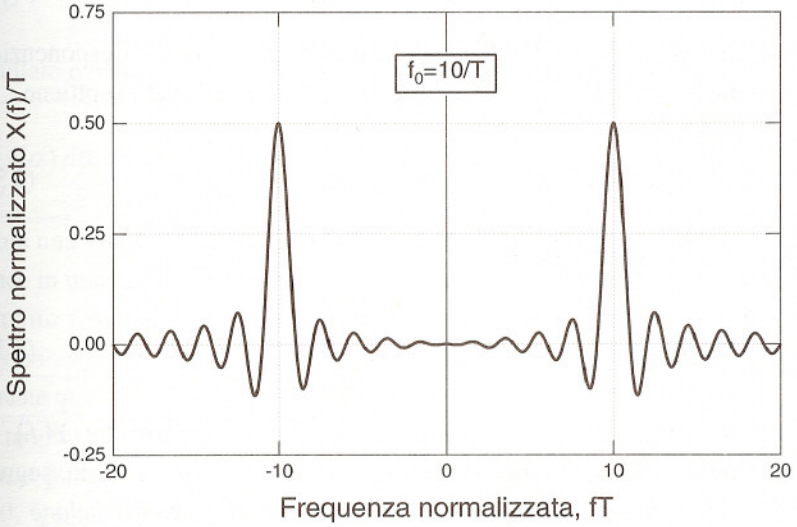


Figura 3.18 Spettro dell'impulso a radiofrequenza

❑

3.3.6 Teoremi di *derivazione e integrazione*

Nella elaborazione dei segnali a tempo continuo, si effettuano spesso operazioni

di derivazione e/o integrazione temporale dei segnali stessi. Sorge quindi la necessità di determinare le trasformate dei nuovi segnali ottenuti con tali operazioni.

Consideriamo dunque come di consueto un segnale $x(t)$ con trasformata $X(f)$. Questi può essere espresso come integrale di Fourier:

$$x(t) = \int_{-\infty}^{\infty} X(f)\, e^{j2\pi ft} df \tag{3.3.28}$$

Se inoltre il segnale è derivabile,

$$\frac{dx(t)}{dt} = \frac{d}{dt} \int_{-\infty}^{\infty} X(f)\, e^{j2\pi ft} df \tag{3.3.29}$$

Procediamo nel calcolo del secondo membro della (3.3.29) invertendo le operazioni di derivazione e di integrazione:

$$\frac{d}{dt} \int_{-\infty}^{\infty} X(f)\, e^{j2\pi ft} df = \int_{-\infty}^{\infty} \frac{d}{dt} \left[X(f)\, e^{j2\pi ft} \right] df = \int_{-\infty}^{\infty} X(f) \frac{d}{dt} e^{j2\pi ft} df \tag{3.3.30}$$

Nell'ultimo passaggio della (3.3.30) è stato sfruttato il fatto che l'esponenziale è l'unica funzione che dipende da t. Calcolando quindi la derivata si ottiene

$$\frac{dx(t)}{dt} = \int_{-\infty}^{\infty} X(f)(j2\pi f)\, e^{j2\pi ft} df \tag{3.3.31}$$

Ponendo $Y(f) = (j2\pi f)X(f)$ nella (3.3.31) si ha infine

$$\frac{dx(t)}{dt} = \int_{-\infty}^{\infty} Y(f)\, e^{j2\pi ft} df \tag{3.3.32}$$

e quindi (si confronti con la (3.3.28)) possiamo affermare che $Y(f)$ è la trasformata della funzione $dx(t)/dt$. Concludiamo allora con la seguente relazione di corrispondenza che prende il nome di *teorema di derivazione*

$$\frac{dx(t)}{dt} \Leftrightarrow j2\pi f \cdot X(f) \tag{3.3.33}$$

L'operazione di derivata temporale di un segnale si traduce, nel dominio della frequenza, in una semplice operazione *algebrica*, e cioè in una *alterazione* di tutte le componenti frequenziali secondo un fattore $j2\pi f$ proporzionale al valore

della frequenza stessa. Oltre a uno sfasamento di $\pm\pi/2$ (a seconda del segno di f), l'operazione di derivata comporta in particolare una *esaltazione delle componenti alle alte frequenze*. Questo fenomeno è evidente dalla Figura 3.19, che mostra la derivata $y(t)$ del segnale $x(t) = \exp(-|t|/T)$ (esponenziale bilatero), lo spettro di tale segnale, e lo spettro di ampiezza $|Y(f)|$ del segnale derivato $dx(t)/dt$.

Risolviamo adesso il problema inverso; indichiamo con $y(t)$ la *funzione integrale* (o segnale integrale, o *primitiva*) di $x(t)$, definita come segue:

$$y(t) = \int_{-\infty}^{t} x(\alpha)\, d\alpha \tag{3.3.34}$$

Poiché ovviamente $x(t) = dy(t)/dt$, utilizzando il teorema di derivazione si ha

$$X(f) = j2\pi f\, Y(f) \tag{3.3.35}$$

da cui si ricava

$$Y(f) = \frac{X(f)}{j2\pi f} \tag{3.3.36}$$

Il risultato è noto come *teorema di integrazione* e si riassume con

$$\int_{-\infty}^{t} x(\alpha)\, d\alpha \Leftrightarrow \frac{X(f)}{j2\pi f} \tag{3.3.37}$$

Ancora una volta, una operazione di calcolo differenziale in ambito temporale si traduce in una semplice operazione *algebrica* (una divisione per il fattore $j2\pi f$) in ambito frequenziale. In questo caso, dualmente al teorema di derivazione, vengono *esaltate le componenti a bassa frequenza* nello spettro del segnale, e attenuate quelle alle alte frequenze.

Il teorema di integrazione è però valido sotto un'ipotesi ben precisa: dal punto di vista puramente algebrico, la (3.3.36) è equivalente alla (3.3.35) solo per $f \neq 0$. Quando $f = 0$, affinché la (3.3.35) possa essere verificata, si deve avere $X(0) = 0$, altrimenti l'uguaglianza è impossibile. La condizione $X(0) = 0$ equivale ad affermare che il segnale $x(t)$ "sottende area nulla". Abbiamo già visto infatti nell' Esempio 3.1 che

$$X(0) = \int_{-\infty}^{\infty} x(t)\, dt \tag{3.3.38}$$

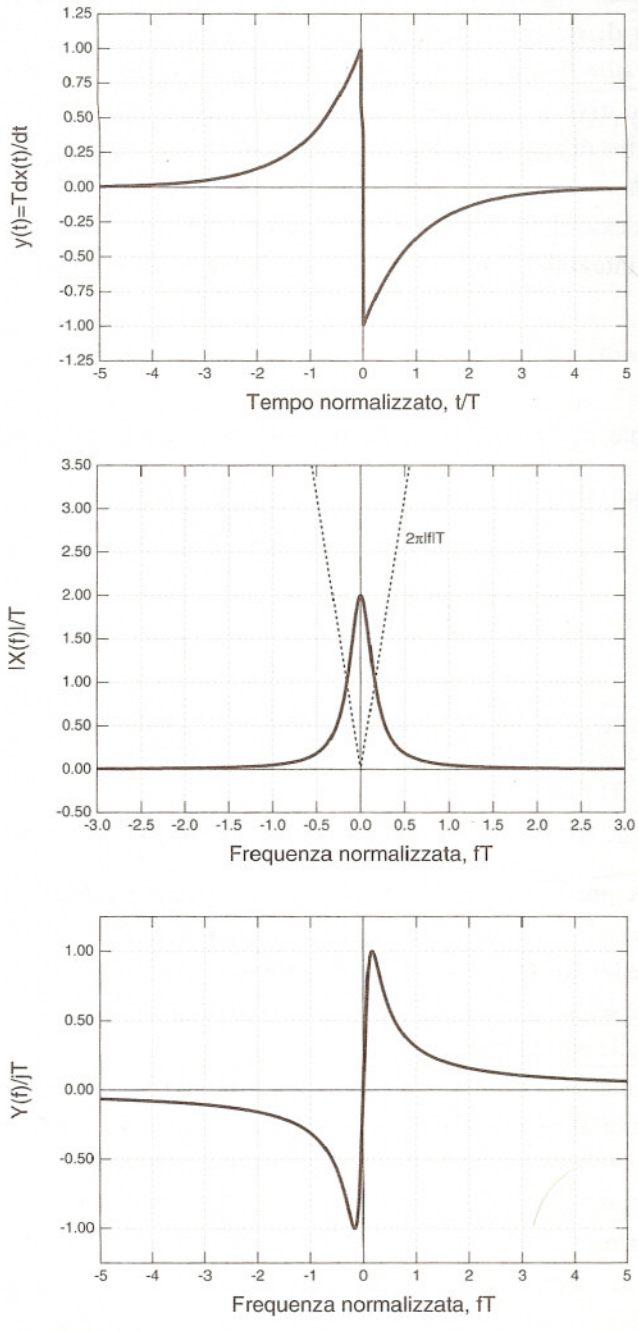


Figura 3.19 Derivata temporale (a) e trasformata (b) del segnale esponenziale bilatero; trasformata della derivata dello stesso segnale (c)

Alternativamente, la stessa condizione può essere espressa dicendo che il segnale $y(t)$ deve tendere a un valore *nullo* quando $t \to \infty$: $y(+\infty) = 0$. Dalla definizione di funzione integrale (3.3.34) si vede infatti che

$$y(+\infty) = \int_{-\infty}^{\infty} x(t)\, dt = X(0) \tag{3.3.39}$$

Un esempio di applicazione del teorema di integrazione è illustrato nella Figura 3.20 in cui il segnale integrato $x(t)$ *sottende area nulla*. Vengono mostrati il segnale integrale $y(t)$ (che ha valore all'infinito nullo), e le due trasformate $X(f)$ e $Y(f)$.

L'ipotesi necessaria per l'applicabilità del teorema di integrazione deriva dal fatto che se la funzione $y(t)$ tendesse a un valore non nullo al tendere di t all'infinito, essa *non avrebbe energia finita* e in generale potrebbe non esistere la sua trasformata di Fourier. Dopo l'introduzione nel Paragrafo 3.4 delle funzioni generalizzate, saremo in grado di ricavare una versione del teorema di integrazione valida *incondizionatamente*.

Esempio 3.9

Vediamo come i teoremi di derivazione e integrazione risultino fondamentali nell'applicazione dell'analisi di Fourier alla teoria dei circuiti lineari.

Consideriamo dunque il circuito elettrico di Figura 3.21 (squadra C-R). Supponendo di conoscere la tensione in ingresso $v_i(t)$ ci proponiamo di calcolare quella in uscita $v_u(t)$. Stavolta, contrariamente a quanto visto nel Paragrafo 2.1, non faremo nessuna ipotesi sulla periodicità del segnale d'ingresso. Le due grandezze sono legate dalla semplice relazione

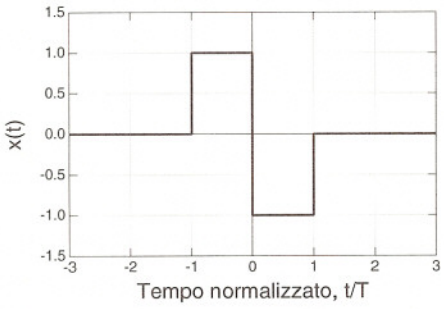
$$v_u(t) = v_i(t) - v_c(t) \tag{E3.9.1}$$

dove $v_c(t)$ è la tensione ai capi del condensatore C,

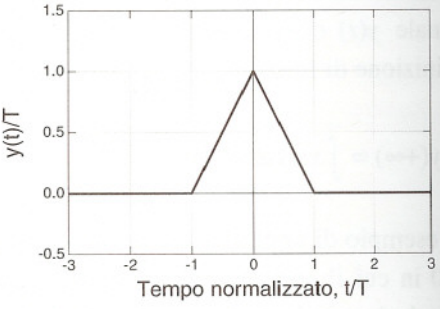
$$v_c(t) = \frac{1}{C} \int_{-\infty}^{t} i(\alpha)\, d\alpha \tag{E3.9.2}$$

La corrente di maglia $i(t)$ è inoltre legata alla tensione di uscita dalla relazione $i(t) = v_u(t)/R$. Sostituendo quest'ultima e la (E3.9.2) nella (E3.9.1) si ricava

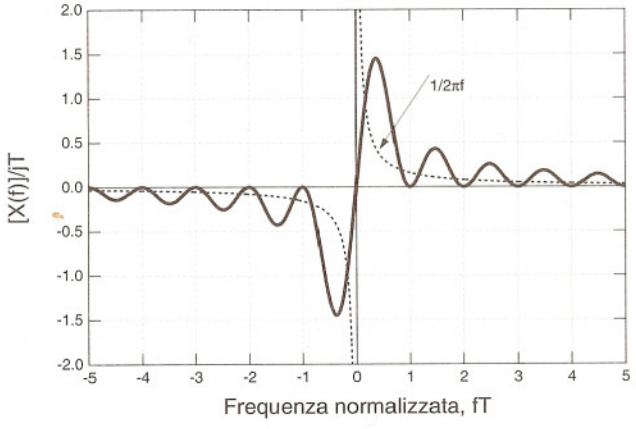
$$v_u(t) = v_i(t) - \frac{1}{C} \int_{-\infty}^{t} \frac{v_u(\alpha)}{R}\, d\alpha \tag{E3.9.3}$$



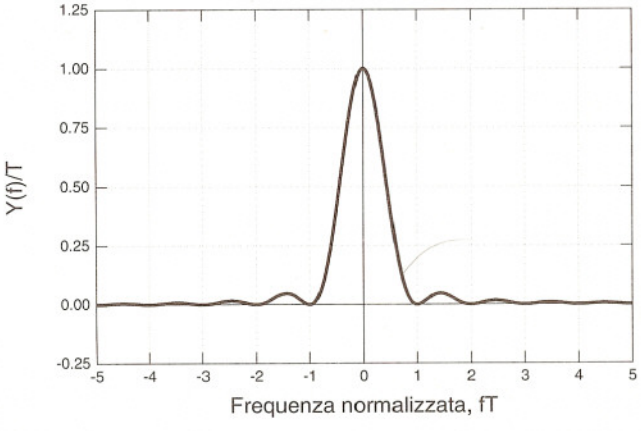
(a)



(b)



(c)



(d)

Figura 3.20 Applicazione del teorema di integrazione: segnale integrando (a), segnale integrale (b), trasformata del segnale integrando (c) e trasformata del segnale integrale (d)

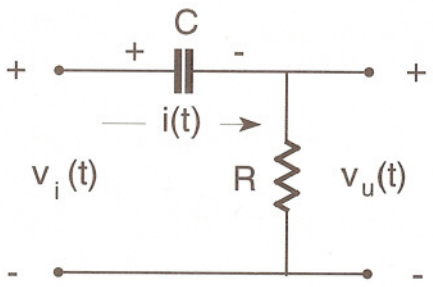


Figura 3.21 Squadra *C-R*

dalla quale, per derivazione, si ottiene l'*equazione differenziale* che caratterizza il comportamento del circuito in ambito temporale:

$$\frac{dv_u(t)}{dt} = \frac{dv_i(t)}{dt} - \frac{1}{C}\frac{v_u(t)}{R} \qquad \text{(E3.9.4)}$$

Poiché il sistema in esame è lineare, è possibile applicare il principio di sovrapposizione degli effetti; questo ci suggerisce di *risolvere il circuito nel dominio della frequenza* pensando di scomporre il segnale come integrale di Fourier. In pratica, questo significa cercare di ricavare la *trasformata* del segnale di uscita in funzione della *trasformata* del segnale d'ingresso.

Supponiamo allora di conoscere $V_i(f)$ e di voler determinare $V_u(f)$. Nelle equazioni da risolvere (E3.9.3) o (E3.9.4) compaiono proprio operazioni di *derivazione o di integrazione* di una funzione. Riconsideriamo ad esempio l'equazione (E3.9.3):

$$v_u(t) = v_i(t) - \frac{1}{C}\int_{-\infty}^{t} \frac{v_u(\alpha)}{R}\, d\alpha \ , \qquad \text{(E3.9.5)}$$

e trasformiamo ambo i membri di questa equazione; sfruttando il teorema di integrazione si ricava:

$$V_u(f) = V_i(f) - \frac{1}{RC}\frac{V_u(f)}{j2\pi f} \quad \Rightarrow \quad V_u(f) = \frac{j2\pi fRC}{1 + j2\pi fRC} V_i(f) \qquad \text{(E3.9.6)}$$

Allo stesso risultato si perviene naturalmente applicando il teorema di derivazione all'equazione differenziale (E3.9.4). Si noti che le trasformate dei segnali di ingresso/uscita $V_i(f)$ e $V_u(f)$ sono legate da una semplice relazione

algebrica e non da una relazione *differenziale* come le corrispondenti grandezze nel dominio del tempo. ❑

Esempio 3.10

Calcoliamo la trasformata continua di Fourier $X(f)$ dell'*impulso triangolare* $x(t)$ di Figura 3.22. Troviamo allora la derivata prima $s(t) = dx(t)/dt$ del segnale dato (ancora rappresentata in Figura 3.22), e osserviamo che essa può essere espressa come segue:

$$s(t) = \frac{A}{T}\,\mathrm{rect}\left(\frac{t+T/2}{T}\right) - \frac{A}{T}\,\mathrm{rect}\left(\frac{t-T/2}{T}\right) \qquad \text{(E3.10.1)}$$

Utilizzando il teorema del ritardo e la trasformata notevole dell'impulso rettangolare (Esempio 3.1) si calcola direttamente la trasformata $S(f)$ di $s(t)$:

$$S(f) = \frac{A}{T}\,T\,\mathrm{sinc}(fT)\,e^{j2\pi fT/2} - \frac{A}{T}\,T\,\mathrm{sinc}(fT)\,e^{-j2\pi fT/2}$$
$$= 2jA\,\mathrm{sinc}(fT)\,\mathrm{sen}(\pi fT) \qquad \text{(E3.10.2)}$$

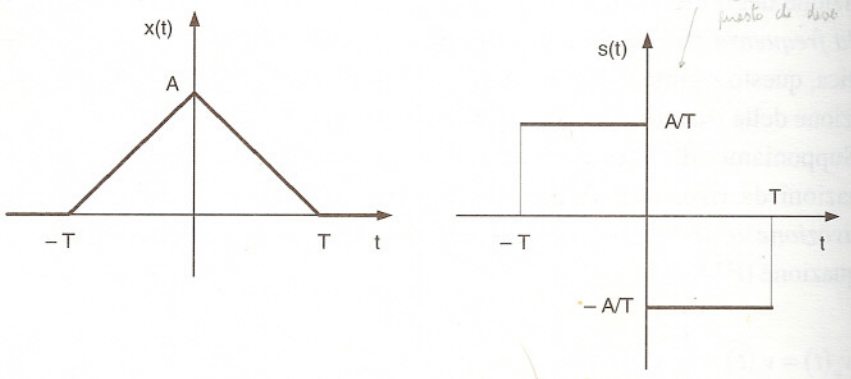


Figura 3.22 Impulso triangolare x(t) e sua derivata s(t)

Osserviamo adesso che $x(t)$ è evidentemente il *segnale integrale* di $s(t)$, e che $x(+\infty) = 0$; è possibile allora utilizzare il teorema di integrazione e ricavare

$$X(f) = \frac{S(f)}{j2\pi f} = \frac{j\,2A\,\mathrm{sinc}(fT)\,\mathrm{sin}(\pi fT)}{j2\pi f} = A\,T\,\mathrm{sinc}^2(fT) \qquad \text{(E3.10.3)}$$

3.3.7 Teorema del *prodotto*

Consideriamo adesso due segnali $x(t)$ e $y(t)$ con le rispettive trasformate di Fourier $X(f)$ e $Y(f)$. Si vuole calcolare la trasformata del segnale *prodotto* $z(t) = x(t)\, y(t)$. Essa è espressa da

$$Z(f) = \int_{t=-\infty}^{+\infty} z(t)\, e^{-j2\pi ft}\, dt = \int_{t=-\infty}^{+\infty} x(t)\, y(t)\, e^{-j2\pi ft}\, dt \tag{3.3.40}$$

Sostituendo a $x(t)$ la sua espressione come *integrale di Fourier* si ricava

$$Z(f) = \int_{t=-\infty}^{+\infty} \left[\int_{v=-\infty}^{+\infty} X(v)\, e^{j2\pi vt}\, dv \right] y(t)\, e^{-j2\pi ft}\, dt \tag{3.3.41}$$

Poiché nel passaggio precedente compaiono due operazioni di integrazione, abbiamo esplicitamente indicato sugli estremi di ogni integrale la variabile di integrazione cui ci si riferisce; inoltre, nell'espressione di $x(t)$ come integrale di Fourier, si è usata una variabile di integrazione "muta" v con un nome differente da f per non creare conflitti con la variabile f di cui la trasformata $Z(f)$ risulta funzione. Se nella (3.3.41) si inverte l'ordine di integrazione, ammettendo che questa operazione sia lecita, si ricava

$$Z(f) = \int_{v=-\infty}^{+\infty} X(v) \left[\int_{t=-\infty}^{+\infty} y(t)\, e^{-j2\pi(f-v)t}\, dt \right] dv \tag{3.3.42}$$

L'integrale entro parentesi quadra rappresenta la trasformata di $y(t)$ calcolata per il valore della frequenza pari a $(f - v)$; di conseguenza la (3.3.42) può essere riscritta come

$$Z(f) = \int_{v=-\infty}^{+\infty} X(v)\, Y(f-v)\, dv = X(f) \otimes Y(f) \tag{3.3.43}$$

L'operazione indicata con il simbolo $\otimes$ prende il nome di *integrale* (o talvolta impropriamente *prodotto*) *di convoluzione*, o convoluzione *tout-court*. La convoluzione, introdotta qui in ambito frequenziale per il calcolo della trasformata del prodotto di due segnali, ha un'importanza cardinale nella teoria dei sistemi lineari stazionari, come vedremo più avanti, nel Capitolo 4.

Dunque il risultato ottenuto può essere riassunto come

$$x(t)\, y(t) \Leftrightarrow X(f) \otimes Y(f) \tag{3.3.44}$$

Si noti che scambiando formalmente il ruolo di $x(t)$ e $y(t)$, per la proprietà *commutativa* del prodotto, anche $Z(f)$ non deve cambiare. Ne segue che l'integrale di convoluzione gode anch'esso della *proprietà commutativa*, nel senso che

$$X(f) \otimes Y(f) = \int_{-\infty}^{+\infty} X(\nu)\, Y(f-\nu)\, d\nu$$

$$= \int_{-\infty}^{+\infty} Y(\nu)\, X(f-\nu)\, d\nu = Y(f) \otimes X(f) \tag{3.3.45}$$

3.3.8 Teorema della *convoluzione*

Consideriamo ora il caso duale del precedente, ovvero un segnale $z(t)$ dato dall'integrale di convoluzione in ambito temporale tra $x(t)$ e $y(t)$:

$$z(t) = x(t) \otimes y(t) = \int_{\alpha=-\infty}^{\infty} x(\alpha)\, y(t-\alpha)\, d\alpha = \int_{\alpha=-\infty}^{\infty} y(\alpha)\, x(t-\alpha)\, d\alpha \tag{3.3.46}$$

e calcoliamone la trasformata di Fourier. Procedendo come nel caso della convoluzione in frequenza si ricava:

$$Z(f) = \int_{t=-\infty}^{+\infty} z(t)\, e^{-j2\pi ft}\, dt = \int_{t=-\infty}^{\infty} \int_{\alpha=-\infty}^{\infty} x(\alpha)\, y(t-\alpha)\, d\alpha\, e^{-j2\pi ft}\, dt$$

$$= \int_{\alpha=-\infty}^{\infty} x(\alpha) \int_{t=-\infty}^{\infty} y(t-\alpha)\, e^{-j2\pi f(t-\alpha+\alpha)}\, dt\, d\alpha$$

$$= \int_{\alpha=-\infty}^{\infty} x(\alpha) \int_{t=-\infty}^{\infty} y(t-\alpha)\, e^{-j2\pi f(t-\alpha)}\, dt\, e^{-j2\pi f\alpha}\, d\alpha$$

$$= \int_{\alpha=-\infty}^{\infty} x(\alpha)\, Y(f)\, e^{-j2\pi f\alpha}\, d\alpha = X(f)\, Y(f) \tag{3.3.47}$$

e quindi

$$x(t) \otimes y(t) \Leftrightarrow X(f)\, Y(f) \tag{3.3.48}$$

Lasciamo al lettore la verifica delle seguenti proprietà di cui gode l'integrale di convoluzione, che vanno ad aggiungersi alla *proprietà commutativa* già discussa nell'ambito del teorema del prodotto:

- *proprietà associativa*:

$$[x(t) \otimes y(t)] \otimes z(t) = x(t) \otimes [y(t) \otimes z(t)] \tag{3.3.49}$$

- *proprietà distributiva rispetto alla somma*:

$$z(t)\otimes[x(t)+y(t)]=z(t)\otimes x(t)+z(t)\otimes y(t) \tag{3.3.50}$$

Prima di esaminare qualche esempio di applicazione dei teoremi del prodotto e della convoluzione, cerchiamo di discutere le operazioni che portano al calcolo di un integrale di convoluzione. Supponiamo quindi che i segnali $x(t)$ e $y(t)$ siano quelli di Figura 3.23, e cerchiamo di trovare il valore che assume $z(t)=x(t)\otimes y(t)$ per un particolare valore del tempo, diciamo t_0.

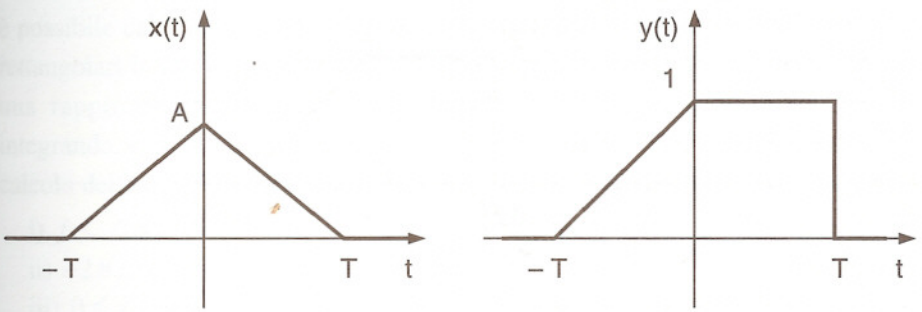


Figura 3.23 Segnali di cui calcolare l'integrale di convoluzione

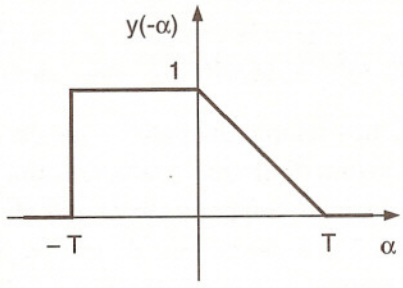
Considerando che

$$z(t_0)=\int_{-\infty}^{\infty}x(\alpha)\cdot y(t_0-\alpha)d\alpha \tag{3.3.51}$$

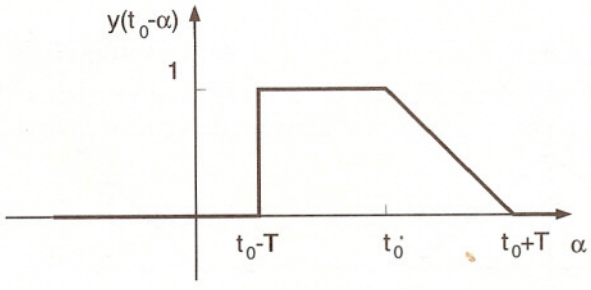
le operazioni idealmente da compiere in questo calcolo sono le seguenti:

- ribaltamento attorno all'asse delle ordinate del segnale $y(\alpha)$ (Figura 3.24a);
- traslazione di $y(-\alpha)$ attorno all'istante t_0 (Figura 3.24b);
- calcolo del prodotto $x(\alpha)\cdot y(t_0-\alpha)$ (Figura 3.24c);
- calcolo dell'integrale del prodotto di cui al passo precedente per ricavare $z(t_0)$ (Figura 3.24d).

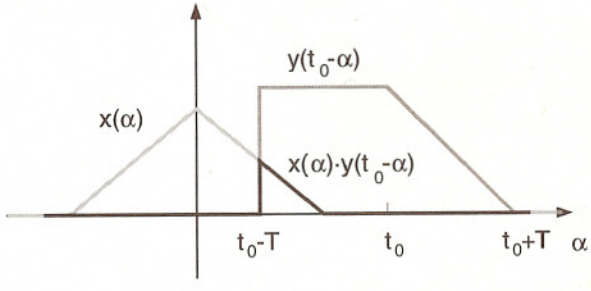
Attraverso questi passi, ripetuti in teoria infinite volte, ovvero per ogni valore dell'istante t_0, si arriva a determinare l'andamento del segnale $z(t)$. In molti casi è semplice procedere alla valutazione dell'integrale di convoluzione "per via grafica", come testimoniano i due esempi seguenti.



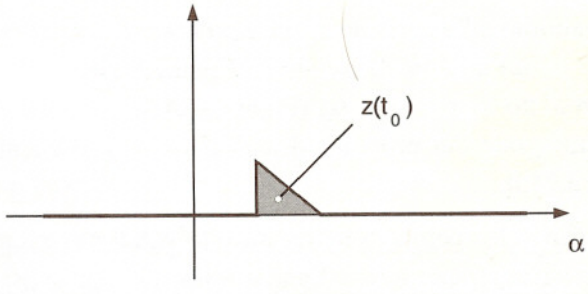
(a)



(b)



(c)



(d)

Figura 3.24 Passi che conducono al calcolo dell'integrale di convoluzione

Esempio 3.11

Come applicazione del teorema del prodotto, consideriamo il caso particolare in cui $x(t) = y(t) = \text{sinc}(2Bt)$; si vuole cioè calcolare la trasformata $Z(f)$ del segnale

$$z(t) = x(t)\, y(t) = \text{sinc}^2(2Bt) \qquad \text{(E3.11.1)}$$

Sapendo che (vedi Esempio 3.3)

$$X(f) = Y(f) = \frac{1}{2B}\text{rect}\left(\frac{f}{2B}\right) \qquad \text{(E3.11.2)}$$

è possibile calcolare $Z(f)$ svolgendo l'integrale di convoluzione di due segnali rettangolari in ambito frequenziale. Come suggerito nel testo, è utile avvalersi di una rappresentazione grafica dei fattori $X(\nu)$ e $Y(f-\nu)$ della funzione integranda, e del loro prodotto $Y(f-\nu)\,X(\nu)$. Dalla Figura 3.25 si nota che il calcolo della convoluzione può essere svolto esaminando i quattro casi seguenti:

i) $f \le -2B$;
ii) $-2B \le f \le 0$;
iii) $0 \le f \le 2B$;
iv) $f \ge 2B$.

Nel caso i) il prodotto fra $Y(f-\nu)$ e $X(\nu)$ è identicamente nullo (Figura 3.25a). Pertanto

$$Z(f) = 0 \quad , \quad f \le -2B \qquad \text{(E3.11.3)}$$

Nel caso ii) (Figura 3.25b) il prodotto fra $Y(f-\nu)$ ed $X(\nu)$ è non nullo, ed è rappresentato in Figura 3.25c. Da questa si deduce che il risultato della operazione di integrazione è espresso dall'area del rettangolo di base $\left[(B+f)-(-B)\right] = 2B+f$ e di altezza $(1/2B)^2$. Segue che

$$Z(f) = \frac{1}{2B}\left[1+\frac{f}{2B}\right] \quad , \quad -2B \le f \le 0 \qquad \text{(E3.11.4)}$$

Analogamente, nel caso iii) (rappresentato in Figura 3.25d) si ricava, con considerazioni di carattere grafico,

$$Z(f) = \frac{1}{2B}\left[1-\frac{f}{2B}\right] \quad , \quad 0 \le f \le 2B \qquad \text{(E3.11.5)}$$

Infine, nel caso iv) si ottiene, analogamente al caso i):

$$Z(f)=0 \ , \ f \geq 2B \tag{E3.11.6}$$

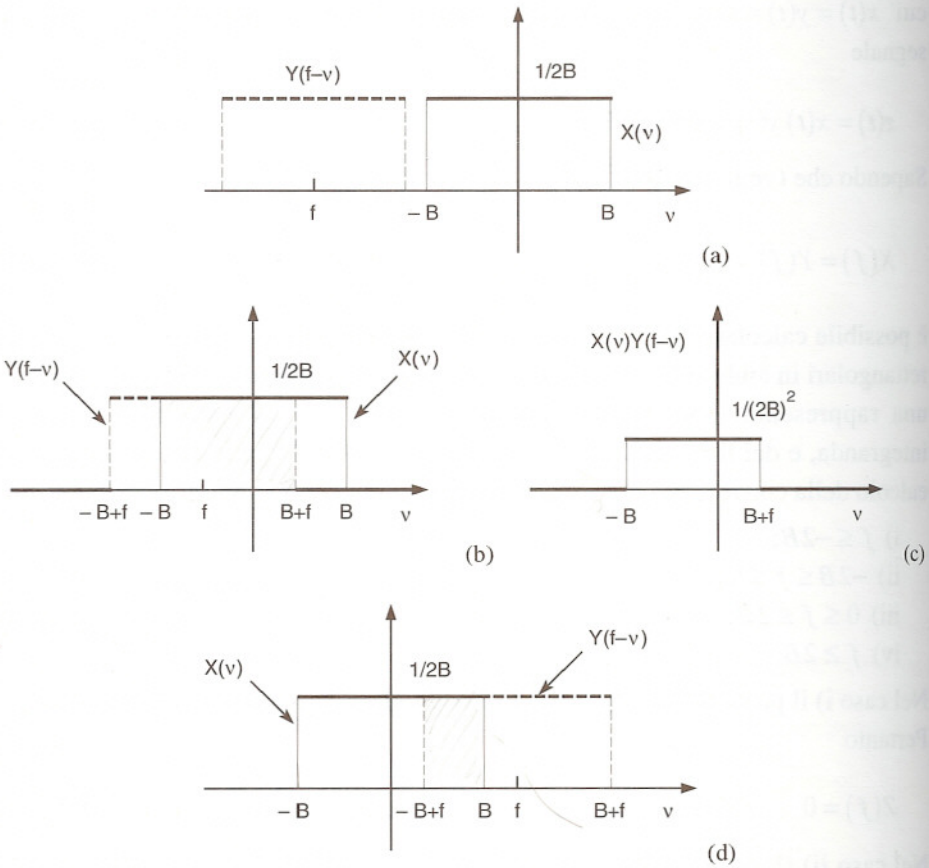


Figura 3.25 Vari casi nel calcolo della convoluzione di un impulso rettangolare con se stesso

Le relazioni (E3.11.3)-(E3.11.6) possono essere riassunte nella seguente espressione:

$$Z(f)=\frac{1}{2B}\left(1-\frac{|f|}{2B}\right)\text{rect}\left(\frac{f}{4B}\right) \tag{E3.11.7}$$

che è rappresentata nella Figura 3.26. Come risultato finale, possiamo scrivere l'ulteriore coppia trasformata-antitrasformata

$$\text{sinc}^2(2Bt) \Leftrightarrow \frac{1}{2B}\left(1-\frac{|f|}{2B}\right)\text{rect}\left(\frac{f}{4B}\right) \tag{E3.11.8}$$

ovvero, ponendo $T = 1/(2B)$,

$$\text{sinc}^2\left(\frac{t}{T}\right) \Leftrightarrow T(1-|f|T)\,\text{rect}\left(\frac{fT}{2}\right) \qquad \text{(E3.11.9)}$$

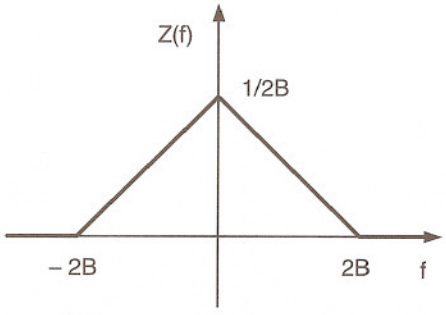


Figura 3.26 Trasformata di Fourier della funzione $\text{sinc}^2(2Bt)$

È importante notare che le funzioni $X(f)$ e $Y(f)$ hanno una "estensione frequenziale" (cioè sono diverse da zero su di un intervallo di ampiezza) pari a $2B$, mentre il loro integrale di convoluzione ha estensione pari a $4B$. In generale l'estensione dell'integrale di convoluzione tra due funzioni aventi estensione limitata è data dalla *somma* delle due estensioni stesse. ❑

Esempio 3.12

Calcoliamo la convoluzione

$$x(t) = s(t) \otimes s(t) \qquad \text{(E3.12.1)}$$

del segnale $s(t) = e^{-t/T}\,\text{u}(t)$ con se stesso. Sappiamo che

$$x(t) = \int_{-\infty}^{\infty} s(\alpha)\, s(t-\alpha)\, d\alpha \qquad \text{(E3.12.2)}$$

L'andamento dei segnali $s(\alpha)$ ed $s(t-\alpha)$ è illustrato in Figura 3.27 (per un caso in cui $t > 0$). Dall'esame di tale figura si vede facilmente che

$$x(t) = 0 \quad , \quad t < 0 \qquad \text{(E3.12.3)}$$

Se invece $t > 0$, e quindi esiste una zona di "sovrapposizione" non nulla tra $s(\alpha)$ ed $s(t-\alpha)$, dalla Figura 3.27 si vede che si ha

$$x(t) = \int_0^t s(\alpha)\, s(t-\alpha)\, d\alpha = \int_0^t e^{-\alpha/T} e^{-(t-\alpha)/T} d\alpha = e^{-t/T} \int_0^t d\alpha = t\, e^{-t/T} \qquad \text{(E3.12.4)}$$

Le (E3.12.3)-(E3.12.4) possono essere riassunte nella

$$x(t) = t\, e^{-t/T}\, \mathrm{u}(t) \qquad \text{(E3.12.5)}$$

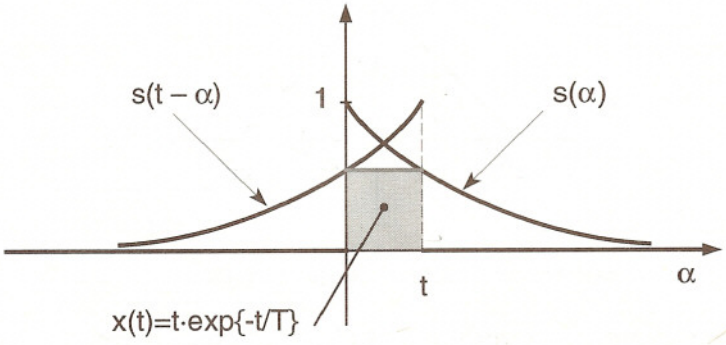


Figura 3.27 Calcolo dell'integrale di convoluzione del segnale $s(t)$ con se stesso

Allo stesso risultato si perviene ovviamente procedendo con il calcolo in maniera puramente analitica. Infatti, sostituendo nella (E3.12.2) l'espressione di $s(t)$, si ricava

$$x(t) = \int_{-\infty}^{\infty} \mathrm{u}(\alpha)\, \mathrm{u}(t-\alpha)\, e^{-\alpha/T} e^{-(t-\alpha)/T} d\alpha = e^{-t/T} \int_{-\infty}^{\infty} \mathrm{u}(\alpha)\, \mathrm{u}(t-\alpha)\, d\alpha \qquad \text{(E3.12.6)}$$

Il lettore dimostri ora che

$$\int_{-\infty}^{\infty} \mathrm{u}(\alpha)\, \mathrm{u}(t-\alpha)\, d\alpha = t\, \mathrm{u}(t) \qquad \text{(E3.12.7)}$$

e che quindi la (E3.12.5) è ancora valida.

In questo esempio abbiamo eseguito la convoluzione tra due segnali *causali*, ottenendo il segnale (E3.12.5) a sua volta causale. Questa proprietà ha validità *generale*, come si può facilmente verificare.

❑

3.4 Trasformate di Fourier *generalizzate*

3.4.1 La funzione generalizzata impulsiva o δ di Dirac

Consideriamo di nuovo la funzione *gradino unitario* (rappresentata in Figura 3.28)

$$u(t)=\begin{cases}1 & t>0\\ 1/2 & t=0\\ 0 & t<0\end{cases} \qquad (3.4.1)$$

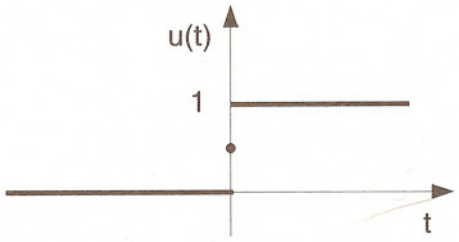


Figura 3.28 Gradino unitario ideale

Evidentemente, il segnale gradino unitario presenta una discontinuità di prima specie per $t=0$. Il valore assegnato alla funzione in questo punto può essere scelto arbitrariamente; per consistenza con il criterio di Dirichlet, si sceglie la semisomma dei limiti destro e sinistro, cioè si pone $u(0)=1/2$.

Immaginiamo di applicare tale segnale in ingresso a un circuito elettrico, ad esempio la squadra *C-R* dell'Esempio 3.9. Il segnale gradino, in questo contesto, è la modellizzazione matematica della "chiusura" del circuito su di un generatore di tensione ideale all'istante $t=0$ di "salita" del gradino stesso, come suggerito dalla Figura 3.29. L'evoluzione temporale dei segnali nel circuito è regolata da un'equazione differenziale simile a quella già ricavata nell'Esempio 3.9. In questo contesto si presenta allora il problema di calcolare la *derivata temporale della funzione gradino* intesa come eccitazione del circuito. Tale operazione presenta delle difficoltà poiché tale derivata è nulla ovunque tranne in $t=0$ ove peraltro essa non è definita, come illustrato in Figura 3.30.

Si può pensare di assegnare alla funzione $d\,u(t)/dt$ un valore arbitrario di comodo in $t=0$, e di usare questa nuova funzione come derivata del gradino unitario. Se questo approccio fosse corretto, si dovrebbe essere in grado di "ricostruire" la $u(t)$ come funzione integrale della sua propria derivata:

$$u(t) = \int_{-\infty}^{t} \frac{d\,u(\alpha)}{d\alpha}\, d\alpha \qquad (3.4.2)$$

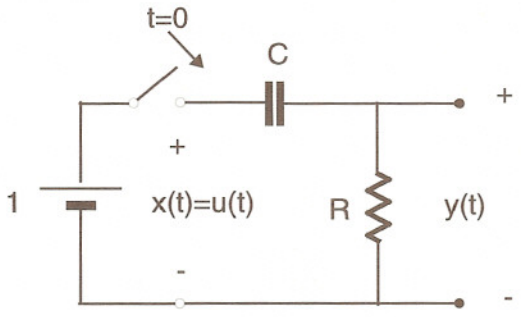


Figura 3.29 Gradino unitario in ingresso a una squadra *C-R*

Tuttavia questo espediente non funziona poiché integrando la *presunta* funzione $du(t)/dt$, *diversa da zero solo in un punto*, non si ottiene la funzione gradino unitario, ma una funzione nulla ovunque.

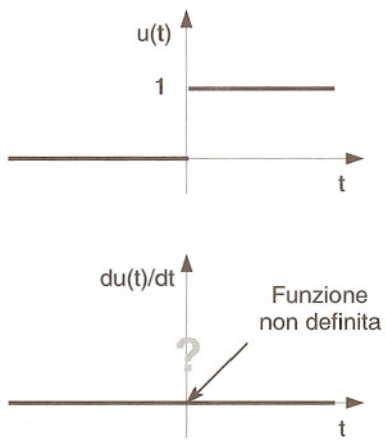


Figura 3.30 Derivata della funzione gradino unitario

Per cercare una soluzione alla questione, osserviamo innanzitutto che la funzione $u(t)$ è comunque un'astrazione matematica poiché i segnali effettivamente presenti nei sistemi fisici non possono presentare discontinuità, cioè salti finiti nel valore del segnale su un tempo nullo. Ogni segnale fisico è caratterizzato da un *tempo di salita* finito, pari al tempo che il segnale "impiega" per passare dal li-

vello 0 prima del "salto" al livello al termine della salita. Il gradino ideale è solo una schematizzazione matematica per quelle situazioni nelle quali la costante di tempo del sistema in cui il segnale compare è molto maggiore del tempo di salita del segnale stesso. Consideriamo allora una migliore approssimazione di una funzione gradino "reale" $u_\varepsilon(t)$ definita dalla relazione:

$$u_\varepsilon(t) = \begin{cases} 0 & t < -\varepsilon \\ (1 + t/\varepsilon)/2 & -\varepsilon \le t \le \varepsilon \\ 1 & t > \varepsilon \end{cases} \tag{3.4.3}$$

illustrata nella Figura 3.31.

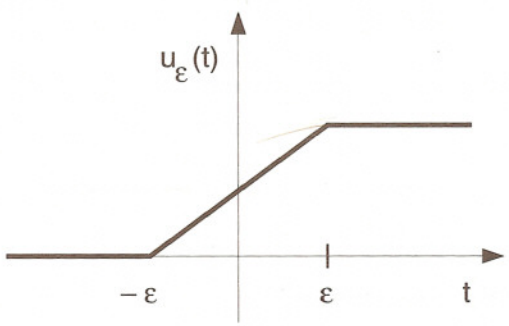


Figura 3.31 Gradino unitario con tempo di salita finito

Nella Figura 3.31 l'andamento del gradino durante l'intervallo di "salita" $(-\varepsilon, \varepsilon)$ è stato preso lineare per semplicità. Le considerazioni che stiamo per fare non cambierebbero però anche se tale andamento fosse ancora più simile a quello di un segnale reale (in particolare ancora più regolare). Il segnale $u_\varepsilon(t)$ è adesso ovunque derivabile, e la sua derivata $\delta_\varepsilon(t) = du_\varepsilon(t)/dt$ è rappresentata in Figura 3.32. Evidentemente,

$$\delta_\varepsilon(t) = \frac{1}{2\varepsilon}\,\mathrm{rect}\left(\frac{t}{2\varepsilon}\right) \tag{3.4.4}$$

e inoltre

$$u_\varepsilon(t) = \int_{-\infty}^{t} \delta_\varepsilon(\alpha)\, d\alpha \tag{3.4.5}$$

senz'alcuna ambiguità.

Osserviamo che riducendo il valore del parametro ε, ossia riducendo il tempo di salita del segnale, si ottiene una approssimazione sempre migliore del gradino ideale, tant'è che si può scrivere

$$u(t) = \lim_{\varepsilon \to 0} u_\varepsilon(t) \tag{3.4.6}$$

e quindi

$$u(t) = \lim_{\varepsilon \to 0} \int_{-\infty}^{t} \delta_\varepsilon(\alpha)\, d\alpha \tag{3.4.7}$$

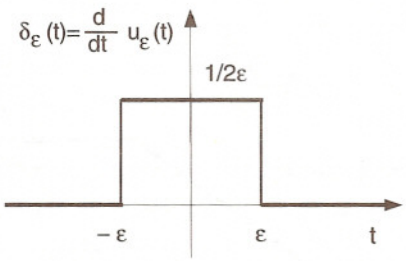


Figura 3.32 Derivata del gradino unitario reale

Il procedimento al limite, perfettamente lecito nell'ambito delle relazioni (3.4.6) e (3.4.7) di cui sopra, diventa però problematico considerando isolatamente il segnale $\delta_\varepsilon(t)$. La Figura 3.33 indica infatti che al tendere di ε a 0 la derivata della funzione gradino "reale" tende ad avere una "durata" sempre più piccola e ad assumere un valore nell'intervallo di salita del gradino sempre più grande. L'aumento di tale valore avviene in maniera inversamente proporzionale a ε poiché dal segnale, per integrazione, si deve sempre ottenere un valore *unitario* pari al salto della funzione gradino. In parole povere, l'*area* degli impulsi rettangolari è comunque unitaria!

Il problema di trovare la "derivata della funzione gradino" sarebbe risolto se nella (3.4.7) potessimo eseguire l'operazione di limite *sotto il segno d'integrale* ponendo dunque

$$\frac{d\,u(t)}{dt} \stackrel{\Delta}{=} \lim_{\varepsilon \to 0} \delta_\varepsilon(t) = \delta(t) \tag{3.4.8}$$

Otterremmo così una funzione $\delta(t)$ con l'apparenza della derivata del gradino ideale come limite della successione di funzioni $\delta_\varepsilon(t)$. Purtroppo, come è

evidente dalla Figura 3.33, il limite di tale successione *non è una funzione* nel senso ordinario dell'analisi matematica. Infatti tale funzione dovrebbe assumere ovunque valore nullo al di fuori del punto $t = 0$ e, se integrata, dovrebbe restituire un valore finito diverso da zero. Non è corretto quindi affermare che $\delta(t)$ è il limite della successione di funzioni $\delta_\varepsilon(t)$. Tuttavia, per estensione, definiamo una *funzione generalizzata impulso unitario* $\delta(t)$ o funzione δ di Dirac, attraverso una sua proprietà di carattere integrale:

$$u(t) = \int_{-\infty}^{t} \delta(\alpha)\, d\alpha \tag{3.4.9}$$

che la identifica come "derivata della funzione gradino", e che è in pratica l'unione della (3.4.7) e della (3.4.8).

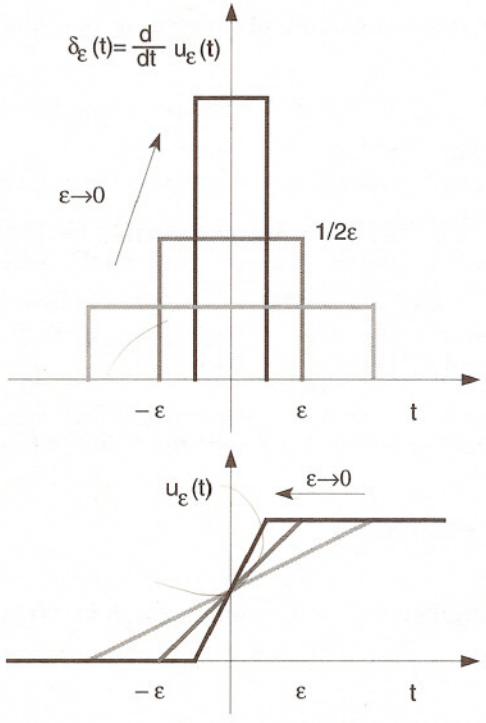


Figura 3.33 Limite del gradino "reale" e della sua derivata

Come abbiamo già discusso, non esiste alcuna funzione ordinaria che possa soddisfare la definizione e quindi dobbiamo ammettere che l'impulso unitario sia un'entità matematica di carattere analogo a quello di una funzione ordinaria,

ma che assume un significato solo quando se ne consideri una qualche proprietà di carattere integrale come nella definizione stessa. Ogniqualvolta si incontra una funzione impulsiva $\delta(t)$ di Dirac, questa deve essere intesa sì come limite della successione $\delta_\varepsilon(t)$, ma non nel senso "diretto" della (3.4.8), privo di significato a rigore; viceversa, l'operazione di limite, per calcolare correttamente il valore dell'integrale in cui la $\delta(t)$ obbligatoriamente compare, deve intendersi eseguita *fuori dal segno di integrale stesso*, nel senso della (3.4.7).

Come esempio di applicazione di questi concetti, consideriamo un segnale $x(t)$ continuo in $t = 0$ e calcoliamo il valore del seguente integrale:

$$I = \int_{-\infty}^{\infty} x(t)\,\delta(t)\,dt \tag{3.4.10}$$

Il problema è ben posto perché la funzione δ compare sotto il segno di integrale; il valore di detto integrale si calcola attraverso un ben inteso procedimento al limite:

$$I = \lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} x(t)\,\delta_\varepsilon(t)\,dt \tag{3.4.11}$$

Come suggerito dalla Figura 3.34, l' integrale della (3.4.11) può essere espresso come

$$\int_{-\infty}^{\infty} x(t)\,\delta_\varepsilon(t)\,dt = \frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} x(t)\,dt = \frac{1}{2\varepsilon}\,2\varepsilon\,x(\tilde{t}) \tag{3.4.12}$$

dove $\tilde{t} \in (-\varepsilon, \varepsilon)$ secondo il *teorema della media*. Il limite (3.4.11) diventa allora

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} x(t)\,\delta_\varepsilon(t)\,dt = \lim_{\varepsilon \to 0} x(\tilde{t}) \tag{3.4.13}$$

Quando $\varepsilon \to 0$, naturalmente $\tilde{t} \to 0$ e quindi, essendo $x(t)$ continua in $t = 0$, si ricava in conclusione

$$\int_{-\infty}^{\infty} x(t)\,\delta(t)\,dt = x(0) \tag{3.4.14}$$

che prende il nome di *proprietà campionatrice dell'impulso unitario* e può essere considerata anche una maniera alternativa di *definire* la funzione $\delta(t)$ attraverso una sua proprietà integrale. Nell'Appendice al Capitolo 3, che può

essere tralasciata in prima lettura, diamo qualche cenno alla *teoria delle distribuzioni* che rappresenta il contesto appropriato per una definizione della $\delta(t)$ di carattere rigoroso, proprio secondo la proprietà campionatrice.

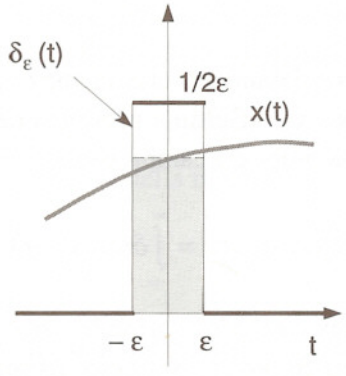


Figura 3.34 Applicazione del teorema della media

Una volta stabilita la maniera generale di intendere e di trattare la funzione generalizzata $\delta(t)$, si può dimostrare che questa gode di molte proprietà (in particolare per quel che riguarda le operazioni di carattere lineare come derivazione, integrazione, moltiplicazione per costante, ecc.) identiche a quelle possedute dalle funzioni *ordinarie*. Queste proprietà autorizzano in pratica a trattare la $\delta(t)$ come un qualunque segnale ordinario a tempo continuo, con le dovute cautele del caso. La Figura 3.35 illustra il simbolo che rappresenta convenzionalmente la funzione $\delta(t)$: il numero posto accanto alla punta della freccia indica che l'area sottesa da questa funzione è unitaria, e *non* è un valore riportato sull'asse delle ordinate.

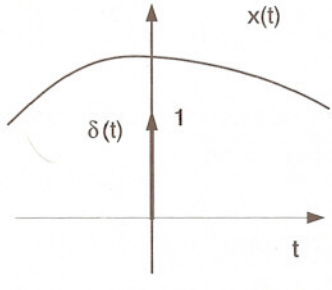


Figura 3.35 Rappresentazione simbolica della funzione δ

3.4.2 Proprietà della funzione generalizzata $\delta(t)$

Abbiamo già esaminato la *proprietà campionatrice* della funzione $\delta(t)$:

$$\int_{-\infty}^{\infty} x(t)\,\delta(t)\,dt = x(0) \tag{3.4.15}$$

Da questa si deduce immediatamente che la funzione $\delta(t)$ è *pari*, cioè che $\delta(t) = \delta(-t)$. Come sempre, quest'ultima proprietà deve essere intesa *in senso integrale* relativamente alla proprietà campionatrice. Infatti possiamo scrivere:

$$\int_{-\infty}^{\infty} \delta(-t)\,x(t)\,dt = -\int_{\infty}^{-\infty} \delta(\alpha)\,x(-\alpha)d\alpha = \int_{-\infty}^{\infty} \delta(\alpha)\,x(-\alpha)\,d\alpha = x(0) \tag{3.4.16}$$

avendo effettuato il cambiamento di variabile $\alpha = -t$. Poiché $\delta(t)$ e $\delta(-t)$ rendono da questo punto di vista lo stesso risultato, possiamo stabilirne l'uguaglianza, da cui segue la proprietà di *parità* appena enunciata.

La proprietà campionatrice può essere estesa considerando la funzione impulsiva "traslata" $\delta(t-t_0)$. Infatti abbiamo

$$\int_{-\infty}^{\infty} x(t)\delta(t-t_0)\;dt = \int_{-\infty}^{\infty} x(t_0+\alpha)\delta(\alpha)\;d\alpha = x(t_0) \tag{3.4.17}$$

Da questa consegue l'ulteriore proprietà

$$x(t)\delta(t-t_0) = x(t_0)\delta(t-t_0) \tag{3.4.18}$$

Questa uguaglianza va intesa ancora nel senso che, dopo una operazione di integrazione, le due funzioni a primo e secondo membro nella (3.4.18) forniscono lo stesso risultato; infatti

$$\int_{-\infty}^{\infty} x(t_0)\,\delta(t-t_0)\,dt = x(t_0)\int_{-\infty}^{\infty} \delta(t-t_0)\,dt = x(t_0) = \int_{-\infty}^{\infty} x(t)\,\delta(t-t_0)\,dt \tag{3.4.19}$$

Elaboriamo adesso la relazione (3.4.17); cambiando variabile di integrazione e tenendo conto che la $\delta(t)$ è pari, si ha

$$\int_{-\infty}^{\infty} x(\alpha)\delta(\alpha-t_0)\;d\alpha = \int_{-\infty}^{\infty} x(\alpha)\delta(t_0-\alpha)\;d\alpha = x(t_0) \tag{3.4.20}$$

e, chiamando t il generico valore t_0, si conclude che

CENTRO BIBLIOTECARIO INGEGNERIA PISA

$$x(t) = \int_{-\infty}^{\infty} x(\alpha)\delta(t-\alpha)\ d\alpha = x(t) \otimes \delta(t) \tag{3.4.21}$$

Questa relazione, piuttosto importante, indica che la funzione generalizzata $\delta(t)$ può intendersi anche, tra le varie interpretazioni, come una sorta di *elemento neutro* rispetto all'operazione di integrale di convoluzione.

Esempio 3.13

Sfruttando formalmente le proprietà della funzione $\delta(t)$, calcoliamo

$$I_{a,b} = \int_{a}^{b} \delta(t)\, dt \tag{E3.13.1}$$

ove a e b sono due parametri reali, $a < b$. È possibile riscrivere questo integrale come

$$I_{a,b} = \int_{-\infty}^{+\infty} \delta(t)\, \mathrm{rect}\left(\frac{t-(b+a)/2}{b-a}\right) dt \tag{E3.13.2}$$

avendo moltiplicato la funzione integranda della (E3.13.1) per una funzione rect(·) identicamente nulla al di fuori dell'intervallo (a,b). Utilizzando la proprietà campionatrice della funzione $\delta(t)$ si ricava allora

$$I_{a,b} = \mathrm{rect}\left(\frac{-(b+a)/2}{b-a}\right) = \mathrm{rect}\left(\frac{1}{2}\cdot\frac{b+a}{b-a}\right) \tag{E3.13.3}$$

Tale risultato può essere semplificato osservando che se

$$|b+a| < b-a \tag{E3.13.4}$$

il modulo dell'argomento della funzione rect(·) nella (E3.13.3) risulta minore di 1/2 e

$$I_{a,b} = 1 \tag{E3.13.5}$$

altrimenti $I_{a,b} = 0$. La condizione (E3.13.4) è verificata solo se a e b hanno *segno discorde* e cioè quando $a < 0$ e $b > 0$.

In pratica, questo risultato si può riassumere dicendo che l'integrale $I_{a,b}$ è diverso da zero (e uguale a 1) se *l'intervallo di integrazione contiene l'origine*, contiene cioè il "punto di applicazione" della funzione $\delta(t)$. Generalizzando il risultato ottenuto si può affermare che

$$\int_a^b x(t)\,\delta(t-t_0)\,dt = x(t_0) \quad \Leftarrow a < t_0 < b \tag{E3.13.6}$$

cioè se il "punto di applicazione" della funzione δ è contenuto nell'intervallo di integrazione (si veda la Figura 3.36a), mentre l'integrale è uguale a zero se la funzione δ è applicata all'esterno dell'intervallo, se cioè $t_0 \notin [a,b]$ (Figura 3.36b).

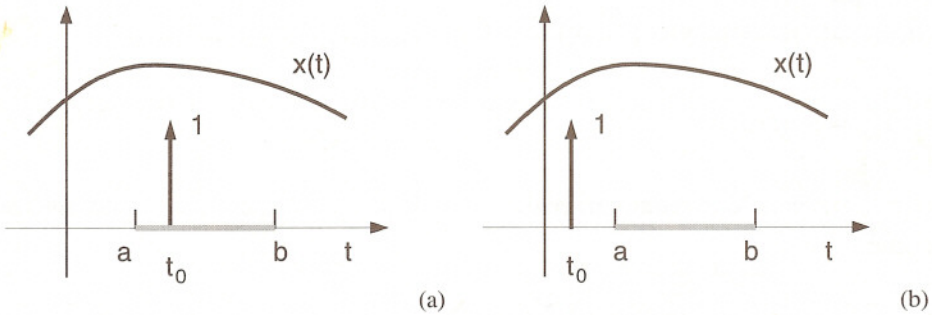


Figura 3.36 Integrali contenenti la funzione δ su intervalli di integrazione limitati

❑

Esempio 3.14

Calcoliamo l'integrale

$$I_a = \int_{-\infty}^{+\infty} x(t)\,\delta(at)\,dt \tag{E3.14.1}$$

ove a è un parametro reale. Effettuando nella (E3.14.1) il cambiamento di variabile $\alpha = at$ si ottiene

$$I_a = \mathrm{sgn}(a)\int_{-\infty}^{+\infty} x\left(\frac{\alpha}{a}\right)\delta(\alpha)\,\frac{d\alpha}{a} \tag{E3.14.2}$$

dove il fattore $\mathrm{sgn}(a)$ mette in conto il fatto che, se $a<0$, il cambiamento di variabile produce una inversione di segno negli estremi di integrazione e quindi fa cambiare il segno del risultato. Dalla (E3.14.2), applicando la proprietà campionatrice della funzione $\delta(t)$, si ricava

$$I_a = \frac{\mathrm{sgn}(a)}{a}\,x(0) = \frac{x(0)}{|a|} \tag{E3.14.3}$$

Avremmo ottenuto lo stesso risultato se nell'integrale originario (E3.14.1) avessimo sostituito alla funzione $\delta(at)$ la funzione $\delta(t)/|a|$. In questo senso abbiamo ricavato un'ulteriore proprietà della funzione impulsiva di Dirac:

$$\delta(at) = \frac{1}{|a|}\delta(t) \qquad \text{(E3.14.4)}$$

❑

3.4.3 Trasformata di Fourier della funzione δ

Calcoliamo ora la trasformata di Fourier della funzione generalizzata $\delta(t)$. Applicando direttamente la definizione di trasformata e ricordando la proprietà campionatrice della $\delta(t)$ si ha

$$\Delta(f) = \int_{-\infty}^{\infty} \delta(t)\, e^{-j2\pi ft}\, dt = e^{-j2\pi ft}\Big|_{t=0} = 1 \qquad (3.4.22)$$

Figura 3.37 Funzione δ di Dirac e relativa trasformata

Quindi lo spettro della funzione δ di Dirac ha ampiezza costante e pari a 1 per ogni valore della frequenza come illustrato in Figura 3.37. La peculiarità della funzione δ è riflessa anche nella peculiarità della sua trasformata: in pratica, lo spettro della funzione δ contiene componenti *a qualunque frequenza* arbitrariamente grande, e tutte con la medesima ampiezza.

La trasformata $\Delta(f)$ del segnale *generalizzato* $\delta(t)$ può anche essere rcavata con un procedimento al limite a partire dalla trasformata $\Delta_\varepsilon(f) = \text{sinc}(2\varepsilon f)$ della funzione *ordinaria* $\delta_\varepsilon(t)$: la Figura 3.38 chiarisce questo punto di vista, e spiega molto bene come deve intendersi lo spettro "piatto" della $\delta(t)$.

Dal teorema di dualità si ricava poi

$$x(t) = 1 \Leftrightarrow \delta(-f) = \delta(f) \qquad (3.4.23)$$

Questo risultato mostra che l'introduzione delle funzioni generalizzate permette di calcolare la trasformata di Fourier di un segnale a *energia infinita* come il segnale costante. Evidentemente, questa trasformata deve intendersi in senso *generalizzato* visto che contiene una funzione generalizzata.

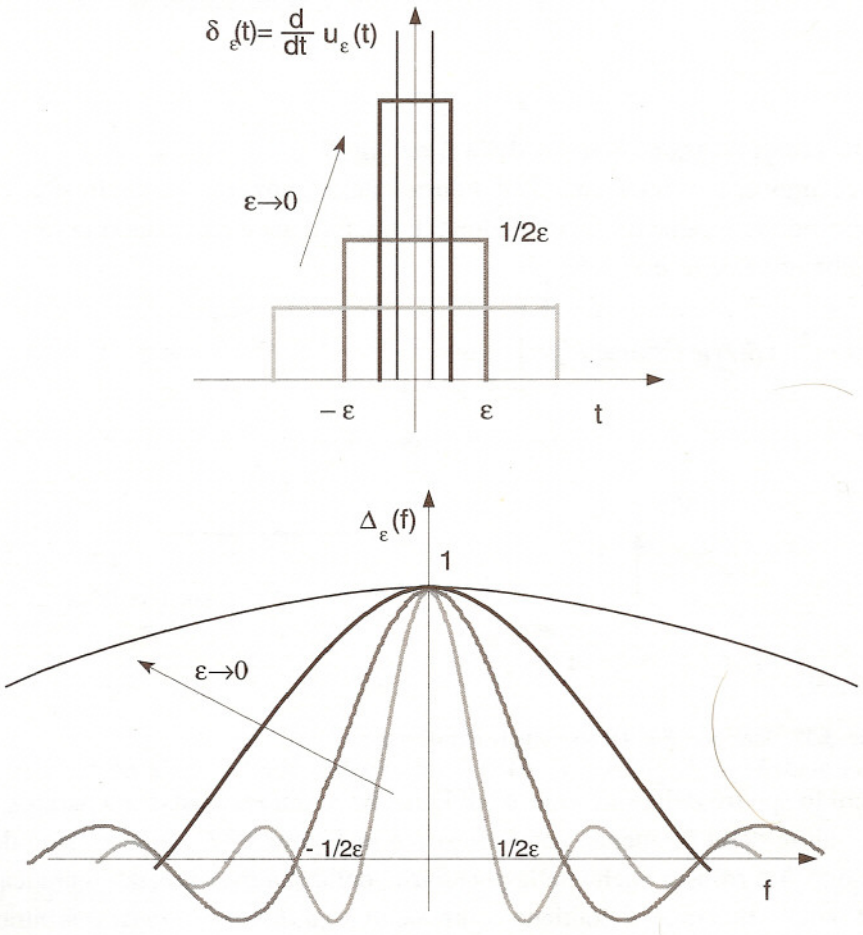


Figura 3.38 Segnale $\delta(t)$ e sua trasformata $\Delta(f)$ intesi come limite

3.4.4 Una trasformata notevole: la funzione 1/*t*

Una trasformata notevole che, come vedremo, è imparentata con le trasformate generalizzate, è quella del segnale $x(t) = 1/t$ il cui grafico è rappresentato in Figura 3.39. Procedendo con il calcolo si ha:

$$X(f)=\int_{-\infty}^{\infty}\frac{e^{-j2\pi ft}}{t}\,dt=\int_{-\infty}^{\infty}\frac{\cos(2\pi ft)-j\sin(2\pi ft)}{t}\,dt=-j\int_{-\infty}^{\infty}\frac{\sin(2\pi ft)}{t}\,dt \quad (3.4.24)$$

ove l'ultimo passaggio si giustifica intuitivamente considerando che la funzione $\cos(2\pi ft)/t$ è dispari e quindi, integrata da $-\infty$ a ∞ rende un risultato nullo. In realtà questa conclusione è lecita solo se dell'integrale generalizzato nella trasformata si considera il cosiddetto *valore principale di Cauchy (VPC)*. A stretto rigore infatti, la funzione $\cos(2\pi ft)/t$ è infinita di ordine 1 nell'origine, e non ammette integrale generalizzato ordinario. Il valore principale di Cauchy invece è quel valore che si ottiene considerando sempre *intervalli di integrazione simmetrici attorno al punto di singolarità* (eventualmente all'infinito), cioè, nel nostro caso:

$$VPC\left[\int_{-\infty}^{\infty}\frac{\cos(2\pi ft)}{t}\right]=\lim_{\substack{\varepsilon\to 0\\ T\to\infty}}\left[\int_{-T}^{-\varepsilon}\frac{\cos(2\pi ft)}{t}+\int_{\varepsilon}^{T}\frac{\cos(2\pi ft)}{t}\right] \quad (3.4.25)$$

che è pari a 0 proprio per la antisimmetria della funzione integranda. Dunque la trasformata cercata può essere anche riscritta come

$$X(f)=-j\int_{-\infty}^{\infty}\frac{\sin(2\pi ft)}{t}\,dt=-j2\pi f\int_{-\infty}^{\infty}\operatorname{sinc}(2ft)\,dt \quad (3.4.26)$$

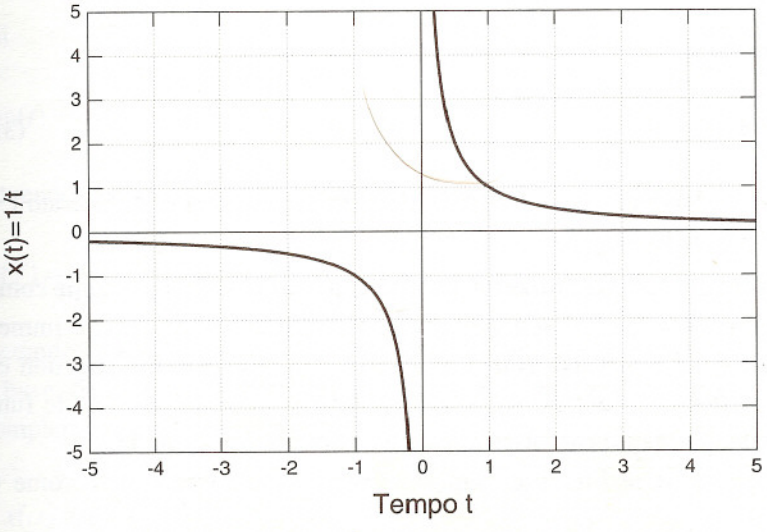


Figura 3.39 Grafico del segnale $x(t)=1/t$

Per il calcolo dell'integrale della (3.4.26) è utile tenere presente la proprietà generale, sfruttata già più volte, per cui l'integrale temporale di un segnale è pari al valore nello zero della relativa trasformata, cioè

$$\int_{-\infty}^{\infty} y(t)\,dt = Y(0) \tag{3.4.27}$$

Richiamiamo allora la consueta trasformata notevole

$$\text{sinc}(2Bt) \Leftrightarrow \frac{1}{2|B|}\text{rect}\left(\frac{f}{2B}\right) \tag{3.4.28}$$

ove compare stavolta a denominatore la quantità $|B|$ per tenere conto del caso $B<0$, caso che non deve alterare l'espressione della trasformata perché la funzione sinc(·) è pari. Da questa trasformata notevole si ricava immediatamente che

$$\int_{-\infty}^{\infty} \text{sinc}(2ft)\,dt = \frac{1}{2|f|}\text{rect}\left(\frac{0}{2f}\right) = \frac{1}{2|f|} \tag{3.4.29}$$

e quindi

$$X(f) = -j2\pi f \int_{-\infty}^{\infty} \text{sinc}(2ft)\,dt = \frac{-j2\pi f}{2|f|} = -j\pi\,\text{sgn}(f) \tag{3.4.30}$$

In conclusione si ricava la trasformata notevole cercata

$$\frac{1}{t} \Leftrightarrow -j\pi\,\text{sgn}(f) \tag{3.4.31}$$

che deve intendersi ancora in senso *generalizzato* per l'aver considerato i valori principali di Cauchy degli integrali coinvolti nella trasformazione.

3.4.5 Trasformata della funzione gradino; teorema d'integrazione completo

Consideriamo di nuovo il segnale gradino unitario ideale u(t); è immediato verificare che la sua trasformata di Fourier in senso *ordinario* non esiste. Cerchiamo allora di stabilire se, alla luce dei risultati ottenuti con le funzioni generalizzate, possiamo calcolare questa trasformata per altra via.

Facendo uso della funzione sgn(t), possiamo esprimere u(t) come (vedi Figura 3.40)

$$u(t) = \frac{1}{2}\text{sgn}(t) + \frac{1}{2} \tag{3.4.32}$$

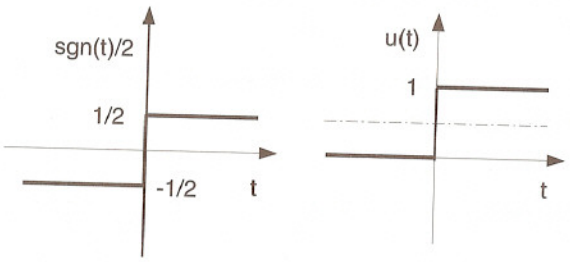


Figura 3.40 Relazione tra le funzioni u(t) e sgn(t)

Calcolando la trasformata di Fourier di entrambi i membri abbiamo

$$U(f) = \frac{1}{2}SGN(f) + \frac{1}{2}\delta(f) \tag{3.4.33}$$

ove $SGN(f)$ rappresenta la trasformata di sgn(t). Ora, $SGN(f)$ può essere facilmente calcolata applicando il teorema di dualità alla trasformata del segnale $1/t$ (3.4.31):

$$-j\pi\, sgn(t) \Leftrightarrow -\frac{1}{f} \tag{3.4.34}$$

da cui

$$\text{sgn}(t) \Leftrightarrow SGN(f) = \frac{1}{j\pi f} \tag{3.4.35}$$

Sostituendo la (3.4.35) nella (3.4.33) si ottiene infine

$$U(f) = \frac{1}{j2\pi f} + \frac{1}{2}\delta(f) \tag{3.4.36}$$

Utilizzando i risultati ottenuti è ora possibile rimuovere l'ipotesi $X(0) = 0$ che è alla base dell'applicabilità del teorema di integrazione nella sua forma "incompleta" ricavata nel Paragrafo 3.3:

$$\int_{-\infty}^{t} x(\alpha)\, d\alpha \Leftrightarrow \frac{X(f)}{j2\pi f} \tag{3.4.37}$$

Infatti, per definizione di integrale di convoluzione, si può scrivere che

$$x(t)\otimes u(t)=\int_{-\infty}^{\infty}x(\alpha)\,u(t-\alpha)\,d\alpha=\int_{-\infty}^{t}x(\alpha)\,d\alpha \tag{3.4.38}$$

Ricordando poi che la trasformata di un prodotto di convoluzione è uguale al prodotto delle trasformate dei fattori, abbiamo

$$\int_{-\infty}^{t}x(\alpha)\,d\alpha=x(t)\otimes u(t)\Leftrightarrow X(f)\,U(f)=X(f)\left[\frac{1}{j2\pi f}+\frac{1}{2}\delta(f)\right] \tag{3.4.39}$$

Il teorema completo d'integrazione afferma quindi che

$$y(t)=\int_{-\infty}^{t}x(\alpha)d\alpha\Leftrightarrow\frac{X(f)}{j2\pi f}+\frac{\delta(f)}{2}X(0) \tag{3.4.40}$$

Il termine aggiuntivo comprende una funzione generalizzata, e naturalmente scompare nell'ipotesi di applicabilità del teorema incompleto, cioè quando $X(0)=0$. Se invece il segnale $x(t)$ non sottende area nulla, la funzione integrale $y(t)$ non tende a zero quando $t\to\infty$, bensì verso il valore finito $X(0)$. Il secondo termine rende conto allora della "componente continua" (cioè del valore medio diverso da zero) pari a $X(0)/2$ che è presente per questo in $y(t)$.

3.4.6 Trasformata delle funzioni seno, coseno e dei segnali periodici

Dunque la funzione generalizzata $\delta(t)$ permette di calcolare trasformate di Fourier non esistenti in senso ordinario. Altri esempi di questo tipo si possono ottenere applicando i teoremi del *ritardo* e della *traslazione in frequenza* alle trasformate generalizzate già ottenute. Ricordando che $\delta(t)\Leftrightarrow 1$, e che $1\Leftrightarrow\delta(f)$, si ottiene immediatamente

$$\delta(t-t_0)\Leftrightarrow e^{-j2\pi ft_0} \tag{3.4.41}$$

$$e^{j2\pi f_0t}=1\cdot e^{j2\pi f_0t}\Leftrightarrow\delta(f-f_0) \tag{3.4.42}$$

Quest'ultima relazione permette poi di calcolare la trasformata continua di Fourier di un'*oscillazione cosinusoidale*. Infatti si ha (vedi Figura 3.41):

$$x_c(t)=\cos(2\pi f_0t)=\frac{e^{j2\pi f_0t}+e^{-j2\pi f_0t}}{2}\Leftrightarrow\frac{\delta(f-f_0)+\delta(f+f_0)}{2} \tag{3.4.43}$$

e, per una oscillazione sinusoidale,

$$x_s(t)=\sin(2\pi f_0t)$$

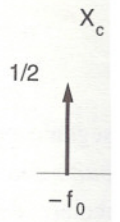


Figura 3.41 Trasformata

Alla luce dei risultat
della modulazione (
prodotto di due segn
scriviamo

$$x(t)\cos(2\pi f_0t)\Leftrightarrow$$

e, poiché

$$X(f)\otimes\delta(f-f_0)=$$

concludiamo, come ci

$$x(t)\cos(2\pi f_0t)\Leftrightarrow$$

e analogamente per la

Avendo determinat
nusoidale in forma rea
continua di un segnale
$x(t)$ periodico con per

$$x_s(t) = \sin(2\pi f_0 t) = \frac{e^{j2\pi f_0 t} - e^{-j2\pi f_0 t}}{2j} \Leftrightarrow \frac{\delta(f - f_0) - \delta(f + f_0)}{2j} \tag{3.4.44}$$

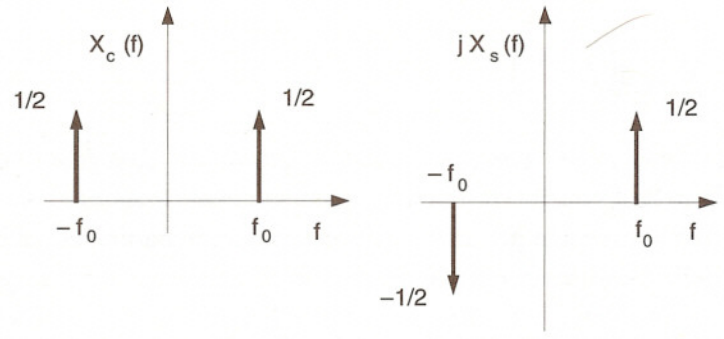


Figura 3.41 Trasformata di Fourier dei segnali $\cos(2\pi f_0 t)$ e $\sin(2\pi f_0 t)$

Alla luce dei risultati ottenuti si può dare una nuova interpretazione al teorema della modulazione (3.3.21); tenendo presente che la trasformata di Fourier del prodotto di due segnali è data dalla convoluzione delle trasformate dei fattori, scriviamo

$$x(t)\cos(2\pi f_0 t) \Leftrightarrow X(f) \otimes \left[\frac{\delta(f - f_0) + \delta(f + f_0)}{2}\right] \tag{3.4.45}$$

e, poiché

$$X(f) \otimes \delta(f - f_0) = \int_{-\infty}^{\infty} X(\alpha)\,\delta(f - f_0 - \alpha)\,d\alpha = X(f - f_0) \tag{3.4.46}$$

concludiamo, come ci aspettavamo, affermando che

$$x(t)\cos(2\pi f_0 t) \Leftrightarrow \frac{X(f - f_0) + X(f + f_0)}{2} \tag{3.4.47}$$

e analogamente per la versione del teorema con la funzione seno.

Avendo determinato la trasformata continua di Fourier di una oscillazione sinusoidale in forma reale o complessa, si riesce anche a esprimere la trasformata continua di un segnale *periodico* qualunque. Ricordiamo infatti che un segnale $x(t)$ periodico con periodo T_0 è sviluppabile in serie di Fourier:

$$x(t) = \sum_{k=-\infty}^{\infty} X_k \, e^{j\frac{2\pi k t}{T_0}} \tag{3.4.48}$$

Calcolando ora la trasformata della serie a *termine a termine* si ricava

$$X(f) = \sum_{k=-\infty}^{\infty} X_k \, \delta\left(f - \frac{k}{T_0}\right) \tag{3.4.49}$$

Questa relazione mostra che il contenuto spettrale di un segnale periodico è *concentrato* nelle frequenze armoniche, piuttosto che *distribuito* con continuità su tutte le frequenze come per un segnale aperiodico; in particolare, il contributo al segnale della k-esima armonica è rappresentato da una funzione δ posizionata in corrispondenza della frequenza k/T_0 e di integrale ("area") pari a X_k, come viene indicato nella Figura 3.42. Lo spettro del segnale periodico ha ancora il tipico andamento "a righe", ma il significato della rappresentazione in Figura 3.42 è diverso da quello degli spettri a righe visti, ad esempio, in Figura 2.4. Infatti nel capitolo precedente le "righe" erano semplicemente proporzionali all'ampiezza del coefficiente di Fourier relativo; in Figura 3.42, viceversa, le "frecce" sono soltanto *rappresentazioni simboliche* della presenza di una funzione δ centrata su quella certa frequenza e nulla più. Anche l'altezza digradante delle funzioni δ indicata in Figura 3.42 è impropria e in un certo senso fuorviante, anche se può dare un'idea dell'andamento dello spettro di ampiezza del segnale.

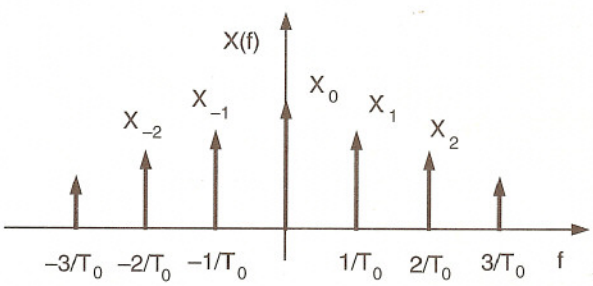


Figura 3.42 Trasformata continua di un segnale periodico con coefficienti di Fourier X_k

Esempio 3.15

Calcoliamo la trasformata del segnale

$$x(t) = \cos^2\left(2\pi t / T_0\right) \tag{E3.15.1}$$

Questo segnale può essere riscritto nella forma

$$x(t)=\frac{1}{2}+\frac{1}{2}\cos\left(\frac{4\pi t}{T_0}\right) \tag{E3.15.2}$$

da cui si ricava:

$$X(f)=\frac{1}{2}\delta(f)+\frac{1}{2}\frac{\delta(f-2/T_0)+\delta(f+2/T_0)}{2}$$
$$=\frac{1}{2}\delta(f)+\frac{1}{4}\delta\left(f-\frac{2}{T_0}\right)+\frac{1}{4}\delta\left(f+\frac{2}{T_0}\right) \tag{E3.15.3}$$

rappresentata in Figura 3.43.

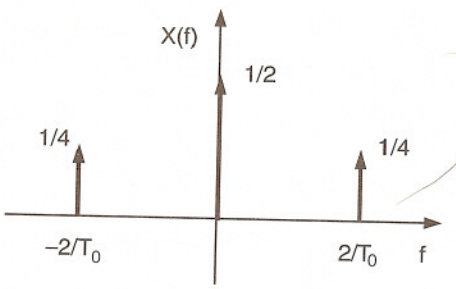


Figura 3.43 Spettro del segnale $x(t)=\cos^2(2\pi/T_0)$ (Esempio 3.15)

❑

Esempio 3.16

Calcoliamo la trasformata $X(f)$ del segnale

$$x(t)=\frac{|t|}{T}\text{rect}\left(\frac{t}{2T}\right) \tag{E3.16.1}$$

rappresentato in Figura 3.44a.

Poiché $x(t)$ è un segnale *lineare a tratti* possiamo utilizzare il teorema dell'integrale per il calcolo di $X(f)$. Ricaviamo dunque la derivata prima $s(t)$ del segnale $x(t)$ (vedi Figura 3.44b)

$$s(t)=\delta(t+T)-\delta(t-T)-\frac{1}{T}\text{rect}\left(\frac{t+T/2}{T}\right)+\frac{1}{T}\text{rect}\left(\frac{t-T/2}{T}\right) \tag{E3.16.2}$$

e la sua trasformata $S(f)$

$$S(f) = e^{j2\pi fT} - e^{-j2\pi fT} - \frac{1}{T}\,T\,\mathrm{sinc}(fT)\,e^{j\pi fT} + \frac{1}{T}\,T\,\mathrm{sinc}(fT)\,e^{-j\pi fT}$$

$$= 2j\,\sin(2\pi fT) - 2j\,\mathrm{sinc}(fT)\,\sin(\pi fT) \qquad \text{(E3.16.3)}$$

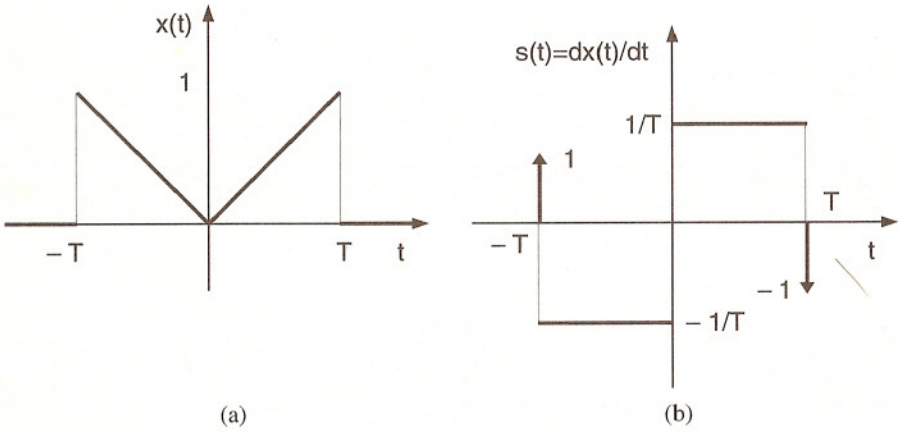


Figura 3.44 Segnali $x(t)$ (a) ed $s(t)$ (b) dell'Esempio 3.16

Poiché $x(+\infty) = 0$ si può applicare il teorema d'integrazione incompleto:

$$X(f) = \frac{S(f)}{j2\pi f} = \frac{2j\,\sin(2\pi fT) - 2j\,\mathrm{sinc}(fT)\,\sin(\pi fT)}{j2\pi f}$$

$$= 2T\,\mathrm{sinc}(2fT) - T\,\mathrm{sinc}^2(fT) \qquad \text{(E3.16.4)}$$

Si ricava lo stesso risultato più rapidamente osservando che il segnale $x(t)$ (E3.16.1) può essere espresso come la *differenza* fra un impulso rettangolare e uno triangolare aventi la stessa durata $2T$:

$$x(t) = \mathrm{rect}\left(\frac{t}{2T}\right) - \left(1 - \frac{|t|}{T}\right)\mathrm{rect}\left(\frac{t}{2T}\right) \qquad \text{(E3.16.5)}$$

dalla quale si deduce immediatamente il risultato (E3.16.4). ❑

Esempio 3.17

Il segnale *periodico* di periodo T_0

$$x(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT_0) \qquad \text{(E3.17.1)}$$

rappresentato in Figura 3.45a è noto come *pettine di* δ. Il coefficiente X_k del suo *sviluppo in serie di Fourier* è

$$X_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi k f_0 t}\, dt \qquad \text{(E3.17.2)}$$

con $f_0 = 1/T_0$. Sulla base dell'Esempio 3.13 e considerando la Figura 3.45a vediamo che l'unico termine del pettine di δ che produce un contributo non nullo nell'operazione di integrazione della (E3.17.2) è quello relativo a $n = 0$, cioè alla funzione δ applicata in $t = 0$, poiché *tutte* le altre funzioni δ sono al di fuori dell'intervallo di integrazione $[-T_0/2, T_0/2]$. Allora

$$X_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \delta(t)\, e^{-j2\pi k f_0 t}\, dt = \frac{1}{T_0} \qquad \text{(E3.17.3)}$$

La rappresentazione in serie di Fourier del segnale $x(t)$ è dunque

$$x(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT_0) = \frac{1}{T_0} \sum_{n=-\infty}^{+\infty} e^{j\frac{2\pi n t}{T_0}} \qquad \text{(E3.17.4)}$$

e la corrispondente *trasformata continua* $X(f)$ di $x(t)$ è

$$X(f) = \sum_{k=-\infty}^{+\infty} X_k\, \delta\left(f - \frac{k}{T_0}\right) = \frac{1}{T_0} \sum_{k=-\infty}^{+\infty} \delta\left(f - \frac{k}{T_0}\right) \qquad \text{(E3.17.5)}$$

Pertanto lo *spettro del segnale pettine di* δ *è anch'esso un pettine di* δ, come illustrato in Figura 3.45b. ❑

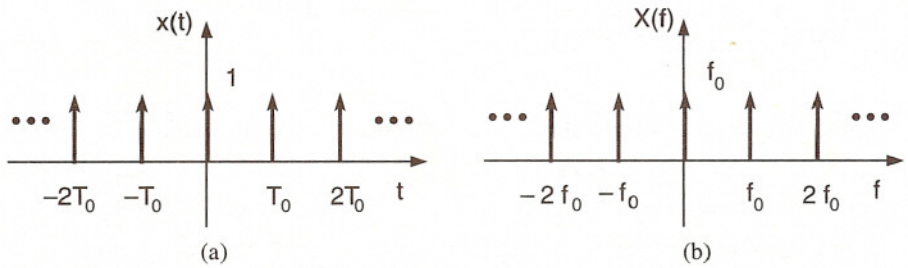


Figura 3.45 Segnale pettine di δ (a) e sua trasformata (b)

3.5 Periodicizzazione e formule di somma di Poisson

Consideriamo un segnale aperiodico $x(t)$ e costruiamo il segnale $y(t)$ periodico di periodo T_0 secondo la relazione di *periodicizzazione*

$$y(t) = \sum_{n=-\infty}^{+\infty} x(t - nT_0) \tag{3.5.1}$$

già incontrata più volte nel Capitolo 2. Il segnale $y(t)$ può essere sviluppato in serie di Fourier:

$$y(t) = \sum_{k=-\infty}^{+\infty} Y_k \, e^{j2\pi k f_0 t} \tag{3.5.2}$$

ove $f_0 = 1/T_0$ e

$$Y_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} y(t) \, e^{-j2\pi k f_0 t} dt \tag{3.5.3}$$

Vediamo come stabilire una relazione fra il *coefficiente* Y_k dello sviluppo in serie (3.5.2) del segnale *periodico* $y(t)$ e la *trasformata* $X(f)$ del segnale base *aperiodico* $x(t)$. Sostituiamo dunque la (3.5.1) nella (3.5.3) ottenendo

$$Y_k = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \sum_{n=-\infty}^{+\infty} x(t - nT_0) \, e^{-j2\pi k f_0 t} dt \tag{3.5.4}$$

Scambiando poi l'operazione di sommatoria e di integrazione nella (3.5.4), si ha

$$Y_k = \frac{1}{T_0} \sum_{n=-\infty}^{+\infty} \int_{-T_0/2}^{T_0/2} x(t - nT_0) \, e^{-j2\pi k f_0 t} dt = \frac{1}{T_0} \sum_{n=-\infty}^{+\infty} \int_{-T_0/2-nT_0}^{T_0/2-nT_0} x(\alpha) \, e^{-j2\pi k f_0 (\alpha + nT_0)} d\alpha$$

$$= \frac{1}{T_0} \sum_{n=-\infty}^{+\infty} \int_{-T_0/2-nT_0}^{T_0/2-nT_0} x(\alpha) \, e^{-j2\pi k f_0 \alpha} d\alpha \tag{3.5.5}$$

Si è giunti a questo risultato dopo aver effettuato il cambiamento di variabile $\alpha = t - nT_0$ e avendo osservato che $e^{-j2\pi k n f_0 T_0} = e^{-j2\pi k n} \equiv 1$. La funzione integranda a secondo membro nella (3.5.5) non dipende dall'indice della serie n; tale indice agisce infatti solo sugli *estremi di integrazione*. Ci si rende allora conto facilmente, con l' ausilio della Figura 3.46, che, al variare di n tra $-\infty$ e $+\infty$, gli intervalli di integrazione $(-T_0/2 - nT_0, T_0/2 - nT_0)$ *della stessa funzione integranda* ricoprono tutto l'asse reale senza sovrapposizioni. Pertanto è

possibile semplificare la (3.5.5) come segue:

$$Y_k = \frac{1}{T_0}\sum_{n=-\infty}^{+\infty}\int_{-T_0/2-nT_0}^{T_0/2-nT_0} x(\alpha)\, e^{-j2\pi k f_0 \alpha}\, d\alpha = \frac{1}{T_0}\int_{-\infty}^{+\infty} x(\alpha)\, e^{-j2\pi k f_0 \alpha}\, d\alpha = \frac{1}{T_0} X\left(\frac{k}{T_0}\right) \qquad (3.5.6)$$

che stabilisce la relazione cercata, detta di *campionamento in frequenza*. I coefficienti di Fourier del segnale periodico $y(t)$ sono dunque, a meno del fattore $1/T_0$, i valori della trasformata continua del segnale-base $x(t)$ in corrispondenza delle *frequenze armoniche* kf_0.

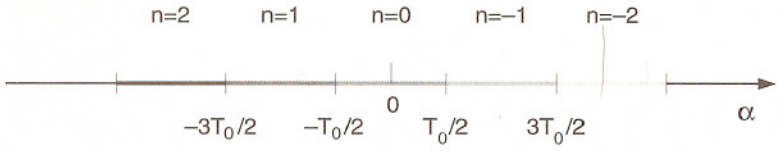


Figura 3.46 Ricopertura dell'asse reale con intervalli di integrazione limitati variabili

Se usiamo l'espressione appena ricavata del coefficiente di Fourier nel corrispondente sviluppo in serie si ottiene

$$\sum_{n=-\infty}^{+\infty} x(t-nT_0) = \sum_{k=-\infty}^{+\infty} \frac{1}{T_0} X\left(\frac{k}{T_0}\right) e^{j\frac{2\pi k t}{T_0}} \qquad (3.5.7)$$

che è nota come *prima formula di somma di Poisson*. Il risultato dell'Esempio 3.17 è un caso particolare di questa formula, con $x(t) = \delta(t)$ e quindi $X(f) = 1$.

Esempio 3.18

Ritroviamo dalla relazione di campionamento in frequenza corrispondente all'operazione di periodicizzazione il valore dei coefficienti di Fourier del segnale cosinusoidale rettificato dell'Esempio 2.5:

$$y(t) = \left|\cos(2\pi f_0 t)\right| \qquad (E3.18.1)$$

È importante notare che l'operazione di "rettificazione" *dimezza* il periodo del segnale. Se infatti il segnale originario $\cos(2\pi f_0 t)$ è periodico di periodo $T_0 = 1/f_0$, dopo la rettifica si ottiene il segnale $y(t)$ periodico di periodo $T_0/2$, come è evidente dalla Figura 3.47. Per usare il campionamento in frequenza, esprimiamo $y(t)$ nella forma

$$y(t) = \sum_{k=-\infty}^{+\infty} x(t - kT_0/2) \tag{E3.18.2}$$

dove abbiamo identificato il "segnale base"

$$x(t) = \cos(2\pi f_0 t)\,\mathrm{rect}\left(\frac{t}{T_0/2}\right) \tag{E3.18.3}$$

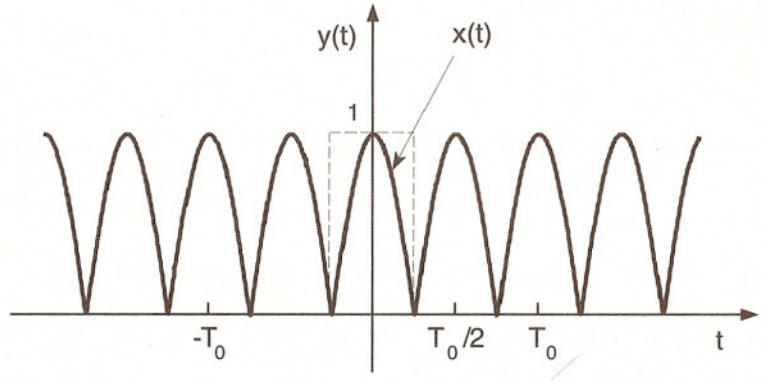


Figura 3.47 Segnale periodico "cosinusoide rettificata" $y(t)$ e suo segnale base $x(t)$

La trasformata continua di Fourier di $x(t)$ si trova facilmente attraverso il teorema della modulazione:

$$X(f) = T_0 \frac{\mathrm{sinc}\big((fT_0 + 1)/2\big) + \mathrm{sinc}\big((fT_0 - 1)/2\big)}{4} \tag{E3.18.4}$$

e il suo andamento è rappresentato in Figura 3.48. Andando a valutare questa trasformata nelle frequenza armoniche $k/(T_0/2)$ come richiesto dalla (3.5.6) si trova immediatamente l'espressione del k-esimo coefficiente di Fourier di $y(t)$:

$$Y_k = \frac{2}{T_0} X\left(\frac{2k}{T_0}\right) = \frac{\mathrm{sinc}(k + 1/2) + \mathrm{sinc}(k - 1/2)}{2} \tag{E3.18.5}$$

che rappresenta proprio il risultato (E2.5.3) (con $A = 1$) dell'Esempio 2.5. ❑

Applichiamo adesso il teorema di dualità alla prima formula di Poisson; si ottiene

$$\sum_{n=-\infty}^{+\infty} X(t-nT_0) = \sum_{k=-\infty}^{+\infty} \frac{1}{T_0} x\left(-\frac{k}{T_0}\right) e^{j\frac{2\pi kt}{T_0}} \Rightarrow \sum_{n=-\infty}^{+\infty} X(t-nT_0) = \sum_{k=-\infty}^{+\infty} \frac{1}{T_0} x\left(\frac{k}{T_0}\right) e^{-j\frac{2\pi kt}{T_0}}$$

$$\Rightarrow \sum_{n=-\infty}^{+\infty} X\left(t-\frac{n}{T}\right) = T \sum_{k=-\infty}^{+\infty} x(kT)\, e^{-j2\pi ktT} \tag{3.5.8}$$

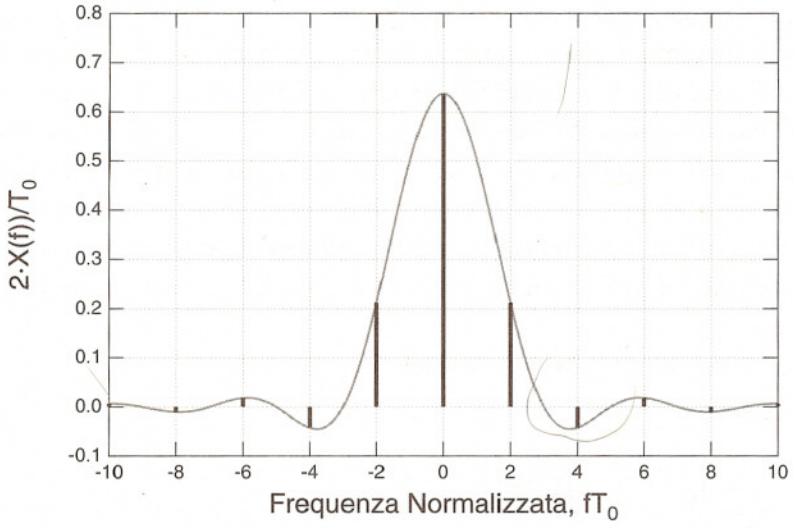


Figura 3.48 Campionamento in frequenza per ottenere i coefficienti di Fourier di $y(t)$

avendo cambiato segno all'indice di sommatoria nel secondo membro e avendo posto $T = 1/T_0$. Se adesso, dal punto di vista puramente formale, cambiamo nome alla variabile corrente da t ad f, otteniamo un'espressione "duale" della (3.5.7) che costituisce la *seconda formula di somma di Poisson*:

$$\sum_{n=-\infty}^{+\infty} x(nT)\, e^{-j2\pi nfT} = \frac{1}{T} \sum_{k=-\infty}^{+\infty} X\left(f-\frac{k}{T}\right) \tag{3.5.9}$$

la cui utilità sarà chiara nel Capitolo 5 a proposito del *campionamento* dei segnali a tempo continuo.

3.6 Relazione fra le trasformate di Laplace e di Fourier

Nel paragrafo precedente abbiamo visto che anche alcuni segnali a potenza finita, come un segnale costante, il segnale gradino o un segnale periodico, ammettono trasformata di Fourier, anche se in senso *generalizzato*. D'altro canto, alcuni segnali possono presentare *potenza illimitata* come il segnale a

rampa:

$$x(t) = t \cdot u(t) = \int_{-\infty}^{t} u(\alpha)\, d\alpha \tag{3.6.1}$$

risultante dall'integrazione del segnale gradino. Questi segnali non ammettono trasformata di Fourier neanche in senso generalizzato. Inoltre, può non essere conveniente trattare i segnali a potenza finita con funzioni generalizzate anche nei casi già visti in cui queste ultime permettono di ricavare la trasformata.

Cerchiamo dunque di risolvere la questione di trovare una funzione *trasformata* di un segnale assegnato che consenta in particolare di trasformare le operazioni differenziali in algebriche, e semplificare quindi le procedure di risoluzione delle equazioni differenziali, come brevemente indicato nell'Esempio 3.9. Limitiamoci inoltre per semplicità al caso dei segnali *causali*, cioè nulli per $t < 0$.

Consideriamo allora il segnale causale per antonomasia, e cioè $x(t) = \mathrm{u}(t)$; sappiamo già che la sua trasformata

$$U(f) = \int_{-\infty}^{\infty} \mathrm{u}(t) e^{-j2\pi ft}\, dt = \int_{0}^{\infty} e^{-j2\pi ft}\, dt \tag{3.6.2}$$

non esiste in senso ordinario, perché l'integrale (generalizzato) appena scritto *non è convergente*. Consideriamo allora un segnale *modificato* $\tilde{x}(t)$ ottenuto da $x(t)$ come illustrato in Figura 3.49, aggiungendo cioè un fattore di "smorzamento" che tenda a far convergere l'integrale (3.6.2):

$$\tilde{x}(t) = x(t) \cdot e^{-\sigma t} \tag{3.6.3}$$

Con l'aggiunta di questo fattore, la funzione integranda tende rapidamente a 0 quando t tende a infinito, e l'integrale soddisfa le condizioni necessarie alla convergenza. Calcolando la trasformata di Fourier del segnale modificato $\tilde{x}(t)$ si ha:

$$\mathcal{F}\left[\tilde{x}(t)\right] = \int_{0}^{\infty} \tilde{x}(t) e^{-j2\pi ft}\, dt = \int_{0}^{\infty} x(t) e^{-\sigma t} e^{-j2\pi ft}\, dt \tag{3.6.4}$$

È chiaro che il valore di σ deve essere scelto opportunamente in relazione all'andamento del particolare segnale. Ad esempio, per il segnale gradino è sufficiente un sia pur minimo grado di smorzamento per far convergere l'integrale, cioè è sufficiente $\sigma > 0$. In tal caso infatti

$$\int_0^\infty x(t)e^{-\sigma t}e^{-j2\pi ft}dt = \int_0^\infty e^{-(\sigma+j2\pi f)t}dt = \frac{e^{-(\sigma+j2\pi f)t}}{\sigma+j2\pi f}\Bigg|_0^\infty = \frac{1}{\sigma+j2\pi f} \tag{3.6.5}$$

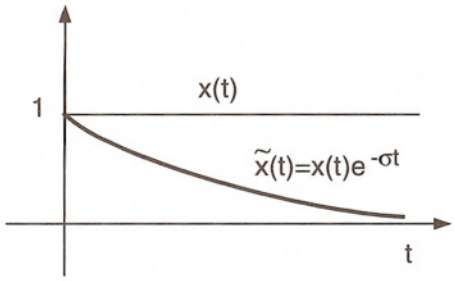


Figura 3.49 Segnale smorzato per il calcolo della trasformata

non appena $\sigma > 0$. Se si considera la quantità $\sigma + j2\pi f = \sigma + j\omega$ come un'*unica variabile complessa*

$$s = \sigma + j\omega = \sigma + j2\pi f \tag{3.6.6}$$

si può interpretare la trasformata di Fourier del segnale "modificato" $\tilde{x}(t)$ (trasformata che verrà a dipendere dal generico valore di σ scelto, oltre che dalla frequenza f) come una *diversa trasformata del segnale originario* $x(t)$ dipendente appunto dalla variabile *complessa* $s = \sigma + j\omega$:

$$X(s) = \int_{-\infty}^{\infty} x(t)e^{-st}dt = \int_0^\infty x(t)e^{-(\sigma+j2\pi f)t}dt = \mathcal{L}[x(t)] \tag{3.6.7}$$

e cioè la *trasformata di Laplace unilatera*[2]. La differenza principale sta evidentemente nel fatto che questa trasformata esiste *in senso ordinario* anche in molti casi in cui la trasformata di Fourier non esiste.

Dalla (3.6.5) è chiaro dunque che il segnale gradino unitario ammette trasformata di Laplace ordinaria e che questa è $U(s) = 1/s$. Altrettanto chiaro dalla breve discussione sul ruolo della "costante di smorzamento" σ è che tale

[2] Applicabile cioè ai soli segnali causali. La trasformata di Laplace *bilatera* è definita mediante un integrale da $-\infty$ a ∞ e si applica a qualunque segnale, analogamente alla trasformata di Fourier. Le considerazioni valide per la trasformata bilatera relativamente alla sua relazione con la trasformata di Fourier sono comunque molto simili a quelle che vengono qui formulate per la unilatera, per cui la trasformata bilatera non verrà ulteriormente presa in considerazione.

risultato deve intendersi valido solo per *particolari* valori di σ, in questo caso $\sigma > 0$. Se ad esempio consideriamo il segnale $x(t) = e^{at}\,u(t)$, $a > 0$, è chiaro che per avere smorzamento sufficiente deve aversi stavolta $\sigma > a$. Relazioni di questo tipo identificano una *zona* del piano complesso della variabile s in cui la trasformata di Laplace esiste, in cui cioè l'integrale nella (3.6.7) è convergente. Questa zona viene chiamata *zona di convergenza* e, per i segnali causali, è un semipiano del tipo $\sigma > \sigma_0$ raffigurato in Figura 3.50; la quantità σ_0 è una caratteristica del segnale $x(t)$ e viene chiamata *ascissa di convergenza*.

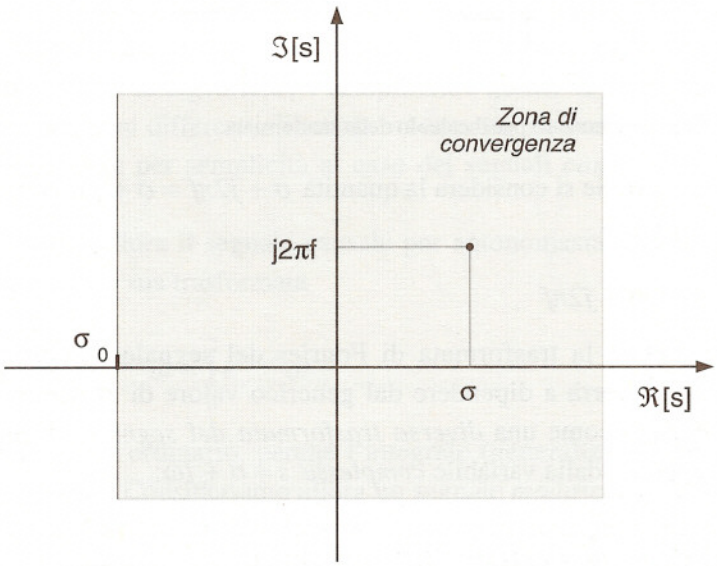


Figura 3.50 Zona di convergenza della trasformata di Laplace sul piano complesso s

Non interessa stabilire qui in dettaglio le caratteristiche della trasformata di Laplace, come la formula di inversione e i vari teoremi e proprietà, molte delle quali sono peraltro analoghe a quelle possedute dalla trasformata di Fourier. È invece interessante approfondire la relazione intercorrente tra le due trasformate. Dal punto di vista formale, se nella (3.6.7) si pone $\sigma = 0$ abbiamo

$$X(s)\big|_{\sigma=0} = \int_0^\infty x(t)\,e^{-st}\,dt\Big|_{\sigma=0} = \int_0^\infty x(t)\,e^{-j2\pi ft}\,dt \tag{3.6.8}$$

che porterebbe la seguente relazione tra le trasformate di Fourier e Laplace per segnali causali:

$$X(f) = X(s)|_{\sigma=0} = X(s)|_{s=j2\pi f} \tag{3.6.9}$$

In pratica, i valori della trasformata di Fourier del segnale sarebbero quelli che la trasformata di Laplace assume per $s = j2\pi f$, cioè sull'*asse immaginario* $\sigma = 0$ del piano s.

Abbiamo usato il condizionale nelle precedenti due affermazioni perché esse sono soggette a una ipotesi. Infatti l'uguaglianza (3.6.9) ha senso e porta a risultati corretti *solo se l'asse immaginario del piano s è contenuto nella zona di convergenza della trasformata di Laplace*. Altrimenti, usare la (3.6.9) può portare a risultati senza senso o errati. La procedura (3.6.9) viene illustrata in Figura 3.51 in un caso permesso; l'andamento della $X(f)$ viene raffigurato su di una "sezione" lungo l'asse immaginario di una superficie che rappresenta i valori di $X(s)$ (supposta reale) in funzione di s.

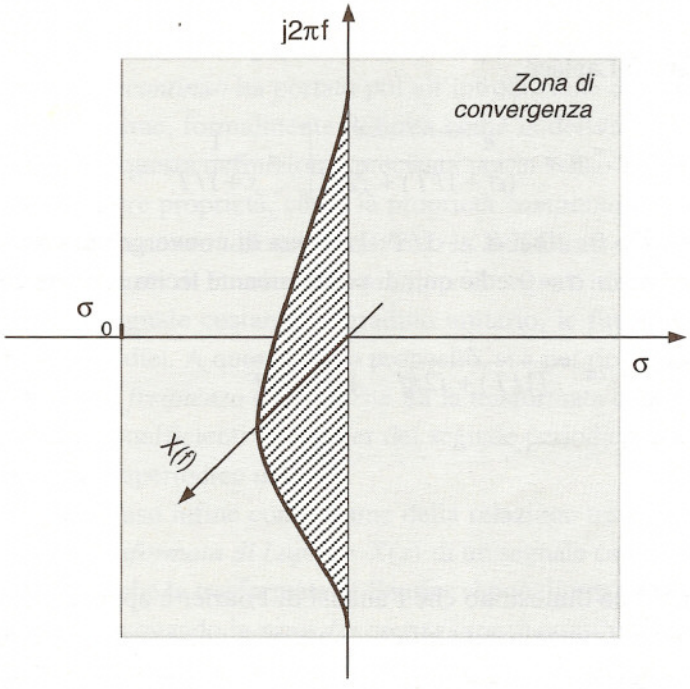


Figura 3.51 Estrazione, quando lecito, della trasformata di Fourier da quella di Laplace

Esempio 3.19

Il segnale gradino $x(t) = \mathrm{u}(t)$ ammette trasformata di Laplace

$$X(s) = \frac{1}{s} \quad , \quad \sigma > 0 \tag{E3.19.1}$$

Applicando la relazione (3.6.9) per tentare di ricavare la trasformata di Fourier del gradino, si ottiene

$$X(s)\big|_{s=j2\pi f} = \frac{1}{j2\pi f} \tag{E3.19.2}$$

che *non* è il risultato corretto. Infatti la zona di convergenza $\Re[s] = \sigma > 0$ della trasformata di Laplace *non* comprende l'asse immaginario $\sigma = 0$. ❑

Esempio 3.20
Il segnale esponenziale unilatero

$$x(t) = e^{-t/T}\,\mathrm{u}(t) \quad , \quad T > 0 \tag{E3.20.1}$$

ha trasformata di Laplace

$$X(s) = \int_0^\infty e^{-t/T} e^{-st}\,dt = -\frac{e^{-[(\sigma+1/T)+j2\pi f]t}}{(\sigma+1/T)+j2\pi f}\Bigg|_0^\infty = \frac{1}{s+1/T} \tag{E3.20.2}$$

purché $\sigma + 1/T > 0$, cioè $\sigma > -1/T$. La zona di convergenza *contiene dunque l'asse immaginario* $\sigma = 0$ ed è quindi perfettamente lecito scrivere che

$$X(f) = X(s)\big|_{s=j2\pi f} = \frac{1}{(1/T)+j2\pi f} = \frac{T}{1+j2\pi fT} \tag{E3.20.3}$$

come già noto. ❑

Sommario

Questo capitolo ha dimostrato che l'analisi di Fourier è applicabile anche ai segnali *aperiodici*. Immaginando infatti di ottenere un segnale aperiodico come il *limite* di un segnale periodico quando il periodo di ripetizione T_0 tende a infinito, si riesce a estendere l'espansione in serie di Fourier, valida per il segnale periodico, anche ai segnali non periodici, ottenendo così l'*integrale* di Fourier. In questa equazione di *sintesi*, il ruolo che nella serie giocavano i coefficienti X_k viene riservato alla *trasformata continua* di Fourier $X(f)$ del segnale aperiodico $x(t)$. Il segnale è ancora scomposto come una *sovrapposizione* di infinite componenti sinusoidali di frequenza variabile *con continuità* da $-\infty$ a $+\infty$, con fase

$\angle X(f)$ e con ampiezza *infinitesima* $|X(f)|\,df$. Molte delle proprietà di simmetria dello spettro di ampiezza $A(f)=|X(f)|$ e dello spettro di fase $\vartheta(f)=\angle X(f)$ del segnale aperiodico sono analoghe a quelle già discusse nel Capitolo 2 (simmetria Hermitiana, segnali pari e dispari ecc.)

Sono state ricavate molte proprietà notevoli della trasformata continua di Fourier, che permettono di calcolare lo spettro di un segnale che subisce particolari operazioni di trasformazione: *ritardo, modulazione, combinazione lineare* tra più segnali, ed è stata messa in evidenza la *dualità* tra i domini del tempo e della frequenza. Tra le proprietà più importanti, menzioniamo i teoremi di *integrazione* e *derivazione*, secondo i quali operazioni di carattere *differenziale* sul segnale temporale equivalgono a più semplici operazioni *algebriche* sulle trasformate. Analoga considerazione può farsi a proposito del teorema della *convoluzione,* per cui l'operazione di integrale di convoluzione nel tempo corrisponde a un semplice *prodotto* in ambito frequenziale.

La necessità di estendere l'operazione di derivata anche in casi in cui il segnale temporale è *discontinuo* ha portato poi all'introduzione della funzione generalizzata $\delta(t)$ di Dirac, formalmente definita come la derivata della funzione gradino unitario. Da questa definizione, precisata poi in senso limite, sono state ricavate numerose altre proprietà, come la proprietà *campionatrice* e la *neutralità* nei confronti della *convoluzione*. Attraverso la δ di Dirac, si è stati in grado di ricavare le trasformate di Fourier *generalizzate* di segnali non trasformabili in senso ordinario: il segnale costante, il gradino unitario, le funzioni seno e coseno, e i segnali periodici. A quest'ultimo proposito, si è poi ricavata la relazione di *campionamento in frequenza* che sussiste tra la trasformata continua di un segnale aperiodico e i coefficienti di Fourier del segnale periodico ottenuto *periodicizzando* il segnale aperiodico dato.

Il capitolo si è chiuso infine con l'esame della relazione tra la trasformata di Fourier $X(f)$ e la *trasformata di Laplace* $X(s)$ di un segnale causale. Si è messo in luce in particolare che la trasformata di Fourier si può direttamente ricavare da quella di Laplace solo quando la *zona di convergenza* di quest'ultima comprende l'asse immaginario $s = j2\pi f$.

Esercizi proposti

3.1 Determinare e rappresentare la trasformata di Fourier $X(f)$ del segnale $x(t)$ nei seguenti tre casi:

a) $x(t) = \begin{cases} \cos^2(\pi t / T) & |t| < T/2 \\ 0 & \text{altrove} \end{cases}$

b) $x(t) = \begin{cases} 0 & t < -T \\ \frac{1}{2}\left(\frac{t}{T}+1\right)^2 & -T \le t < 0 \\ 1-\frac{1}{2}\left(\frac{t}{T}-1\right)^2 & 0 \le t < T \\ 1 & t \ge T \end{cases}$

c) $x(t)$ come in Figura 3.52.

3.2 Dato un segnale $x(t)$ con trasformata $X(f)$, trovare e rappresentare la trasformata del segnale

$$y(t) = x(t) \cdot \sum_n \delta(t - nT)$$

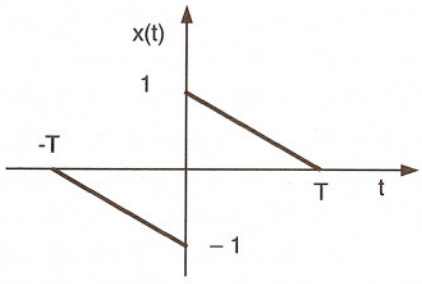


Figura 3.52

3.3 Determinare le trasformate dei seguenti segnali *semiperiodici* (cioe periodici per $t > 0$):

i) $x_c(t) = \cos(2\pi f_0 t) \cdot u(t)$

ii) $x_s(t) = \sin(2\pi f_0 t) \cdot u(t)$

iii) $y(t) = A\cos(2\pi f_0 t + \varphi) \cdot u(t)$

3.4 Determinare la relazione *duale* del teorema di derivazione (calcolare cioè $\mathcal{F}^{-1}[dX(f)/df]$), e sfruttarla per calcolare la trasformata del segnale $x(t)$ in Figura 3.53.

3.5 Calcolare e rappresentare $z(t) = x(t) \otimes y(t)$, con

$$x(t) = e^{-t/T}u(t) \quad , \quad y(t) = \text{rect}\left(\frac{t - T/2}{T}\right)$$

3.6 Rappresentare la seguente trasformata di Fourier:

$$X(f)=\begin{cases} T & |f|\le 1/2T \\ 2T(1-|f|T) & 1/2T<|f|\le 1/T \\ 0 & |f|>1/T \end{cases}$$

e trovare il segnale $x(t)$ antitrasformato.

3.7 Determinare la trasformata di Fourier $D(f)$ del segnale

$$d(t)=\sum_k (-1)^k \delta(t-kT_0/2)$$

Utilizzando questo risultato dimostrare poi che i coefficienti di Fourier di ordine pari di un qualunque segnale periodico alternativo sono nulli.

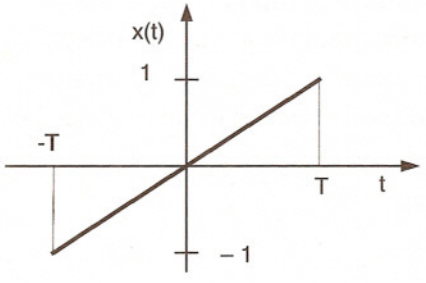


Figura 3.53

3.8 Determinare i coefficienti X_k della serie di Fourier del segnale periodico $x(t)$ di Figura 3.54.

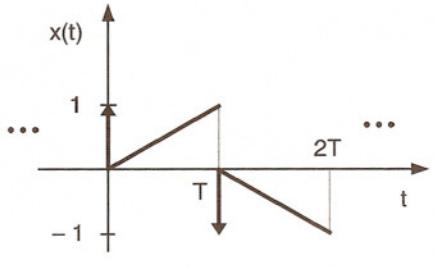


Figura 3.54

3.9 Un segnale periodico $y(t)$ di periodo T_0 è ottenuto come periodicizzazione del segnale aperiodico $x(t)$:

$$y(t) = \sum_{n=-\infty}^{+\infty} x(t - nT_0)$$

Dimostrare che

$$y(t) = x(t) \otimes \sum_{n=-\infty}^{+\infty} \delta(t - nT_0)$$

e da questa espressione ricavare in una maniera alternativa a quella del testo la relazione di *campionamento in frequenza* tra Y_k e $X(f)$.

3.10 Il segnale cosiddetto *coseno rialzato*

$$x(t) = \cos^2\left(\frac{\pi t}{2T}\right) \text{rect}\left(\frac{t}{2T}\right)$$

viene periodicizzato con periodo T. Determinare i coefficienti di Fourier del segnale periodicizzato tramite la formula del campionamento in frequenza, e discutere a posteriori il risultato ottenuto. Era questo prevedibile in partenza? Ripetere lo stesso procedimento per il segnale

$$y(t) = \left(1 - \frac{|t|}{T}\right) \text{rect}\left(\frac{t}{2T}\right)$$

Appendice

Cenni alla teoria delle distribuzioni

A.1 Definizione di distribuzione e di funzione generalizzata

Senza nessuna pretesa di rigore matematico, diamo nelle pagine seguenti alcuni cenni alla teoria delle distribuzioni per chiarire meglio il contesto in cui le funzioni generalizzate hanno una sistemazione formale corretta.

Si definisce *funzionale* e si indica con $T[x]$ una corrispondenza che associa a una funzione $x(t)$, di durata limitata, un ben definito valore numerico reale o complesso. Un esempio di funzionale è l'energia di un segnale:

$$T[x] = E_x = \int_{-\infty}^{\infty} x^2(t)dt \tag{A.1}$$

o il suo limite superiore:

$$T[x] = \sup_t [x(t)] \tag{A.2}$$

o anche l' "area sottesa" dallo stesso

$$T[x] = \int_{-\infty}^{\infty} x(t)dt \tag{A.3}$$

Il funzionale è una generalizzazione del concetto di funzione nel senso di corrispondenza o mappa. Contrariamente alle funzioni ordinarie, gli elementi del dominio della corrispondenza sono *funzioni* (cioè segnali) anziché valori numerici, e il valore reso dal funzionale stesso dipende dall'andamento del

segnale nella sua globalità.

Distinguiamo alcune categorie di funzionali. Un funzionale è *lineare* se per esso vale la relazione

$$T[\alpha\, x_1(t)+\beta\, x_2(t)]=\alpha\, T[x_1(t)]+\beta\, T[x_2(t)] \tag{A.4}$$

per ogni valore dei parametri α e β e per ogni coppia di funzioni $x_1(t)$ e $x_2(t)$. Evidentemente, il funzionale che associa a ogni segnale la propria energia *non* è lineare, mentre quello che associa l' "area sottesa" lo è. Inoltre, un funzionale è *continuo* se, data una successione di funzioni $x_\varepsilon(t)$ con $\lim_{\varepsilon\to 0} x_\varepsilon(t)=x(t)$, è verificata la relazione

$$\lim_{\varepsilon\to 0} T[x_\varepsilon(t)]=T\left[\lim_{\varepsilon\to 0} x_\varepsilon(t)\right]=T[x(t)] \tag{A.5}$$

per ogni valore del tempo t. La continuità garantisce in pratica che è possibile "portare il limite sotto il segno di funzionale". Un funzionale lineare e continuo è una *distribuzione.*

Esempio A.1

Consideriamo il funzionale definito dalla relazione seguente:

$$T[x(t)]=x(0) \tag{A.6}$$

Come possiamo notare, questo particolare funzionale dipende da un solo valore del segnale $x(t)$; verifichiamo se esso gode delle proprietà di linearità e continuità. La (A.6) implica che

$$T[\alpha\, x_1(t)+\beta\, x_2(t)]=\alpha\, x_1(0)+\beta\, x_2(0)=\alpha\, T[x_1(t)]+\beta\, T[x_2(t)] \tag{A.7}$$

che dimostra la linearità del funzionale. Per quanto riguarda la continuità, se è verificata la seguente relazione

$$\lim_{\varepsilon\to 0} x_\varepsilon(t)=x(t) \qquad \forall t \tag{A.8}$$

possiamo scrivere che

$$\lim_{\varepsilon\to 0} T[x_\varepsilon(t)]=\lim_{\varepsilon\to 0} x_\varepsilon(0)=x(0)=T[x(t)] \tag{A.9}$$

e la continuità resta dimostrata. Il funzionale in esame è pertanto una *distribuzione.* ❑

Nello studio della teoria delle distribuzioni è utile introdurre una funzione di comodo $\varphi(t)$ che permetta di esprimere il valore del funzionale nella seguente forma:

$$T_{\varphi}[x] = \int_{-\infty}^{\infty} \varphi(t)\, x(t)\, dt \tag{A.10}$$

Si richiede che la funzione $\varphi(t)$ sia *localmente sommabile*, cioè che verifichi la seguente relazione:

$$\int_{t_1}^{t_2} |\varphi(t)|\, dt < \infty \qquad \forall\, [t_1, t_2] \tag{A.11}$$

con t_1 e t_2 finiti. Il motivo per cui si introduce tale rappresentazione è che essa garantisce automaticamente la linearità e continuità del funzionale. In altre parole, qualunque sia la funzione $\varphi(t)$ scelta, la $T_{\varphi}[x]$ è comunque una *distribuzione*. Se infatti la proprietà di *linearità* è garantita dalla presenza dell'operatore di integrale, la *continuità* del funzionale segue immediatamente dalla locale sommabilità di $\varphi(t)$, come è semplice dimostrare. Il pedice φ nella (A.10) indica che la distribuzione si "appoggia" alla funzione $\varphi(t)$. Se, ad esempio, scegliamo $\varphi(t) = 1$ otteniamo la distribuzione che fornisce il valore dell'area sottesa dalla curva $x(t)$. Invece se $\varphi(t) = u(t_0 - t)$ con t_0 assegnato si ottiene

$$T_{\varphi}[x] = \int_{-\infty}^{\infty} x(t)\, u(t_0 - t)\, dt = \int_{-\infty}^{t_0} x(t)\, dt \tag{A.12}$$

cioè il funzionale restituisce il valore algebrico dell'area sottesa dal segnale $x(t)$ fino all'istante t_0.

A.2 La funzione generalizzata δ di Dirac

Dunque ogni funzione ordinaria $\varphi(t)$ genera una distribuzione, ma viceversa non è detto che data una qualunque distribuzione si riesca a trovare una funzione ordinaria $\varphi(t)$ che la generi. Infatti, se consideriamo la distribuzione $T[x] = x(0)$ ci rendiamo conto che non esiste alcuna funzione ordinaria $\varphi(t)$ che permetta di scrivere tale distribuzione nella forma (A.10). Si definiscono allora le *funzioni generalizzate* che garantiscono la possibilità di scrivere ogni distribuzione nella forma (A.10); tra queste, particolare importanza assume la *funzione impulsiva* δ *di Dirac* implicitamente definita dalla distribuzione appena citata:

$$T_\delta[x] \triangleq \int_{-\infty}^{\infty} x(t)\,\delta(t)\,dt = x(0) \tag{A.13}$$

Questa è la definizione formalmente corretta per la funzione $\delta(t)$; essa prevede ovviamente che la $\delta(t)$ debba comparire sotto il segno di integrale. La (A.13) è gia stata presentata nel testo come una *proprietà* della $\delta(t)$ ricavabile a partire dalla definizione *euristica* intesa come limite di una successione di funzioni ordinarie.

A.3 Derivata di una distribuzione

L'introduzione della $\delta(t)$ nel testo è stata giustificata attraverso la ricerca della derivata temporale del gradino unitario; si è cioè stabilito intuitivamente che

$$\delta(t) = \frac{d\,\mathrm{u}(t)}{dt} \tag{A.14}$$

Cerchiamo adesso di dimostrare la correttezza di questa relazione nell'ambito della teoria delle distribuzioni. Per fare ciò, definiamo innanzitutto la *derivata di una distribuzione* che si appoggia alla funzione $\varphi(t)$ come

$$T'_\varphi[x] \triangleq T_{\varphi'}[x] = \int_{-\infty}^{\infty} \varphi'(t)\,x(t)\,dt \tag{A.15}$$

ove $\varphi'(t)$ indica la derivata prima di $\varphi(t)$ (nell'ipotesi di esistenza). Integrando per parti il secondo membro della (A.15) si ottiene

$$T'_\varphi[x] = \varphi(t)\,x(t)\Big|_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \varphi(t)\,x'(t)\,dt \tag{A.16}$$

Il primo addendo della (A.16) è nullo poiché $x(t)$, per avere energia finita, deve tendere a zero quando il tempo tende all'infinito; allora si può dare un'altra definizione di derivata di una distribuzione:

$$T'_\varphi[x] = -T_\varphi[x'] \tag{A.17}$$

Utilizzando questo risultato cerchiamo di giustificare la relazione (A.14). Ponendo $\varphi(t) = \mathrm{u}(t)$ nella (A.10) si ottiene la distribuzione

$$T_u[x] = \int_0^{\infty} x(t)\,dt \tag{A.18}$$

e, calcolandone la derivata attraverso la (A.17), si ricava

$$T'_u[x] = -\int_{-\infty}^{\infty} u(t)\, x'(t)\, dt = -\int_{0}^{\infty} x'(t)\, dt = \int_{0}^{-\infty} x'(-\alpha)\, d\alpha = x(0) = T_\delta[x] \qquad \text{(A.19)}$$

Dunque, poiché la derivata della distribuzione che si appoggia ad u(t) è la distribuzione che si appoggia a $\delta(t)$, allora, secondo la (A.15), la $\delta(t)$ può essere considerata la *derivata della funzione* $u(t)$.

Definiamo adesso una nuova distribuzione:

$$T_D[x] = -x'(0) \qquad \text{(A.20)}$$

in cui, per così dire, si campiona il valore nell'origine della *derivata* del segnale $x(t)$ cambiata di segno. Chiamiamo $D(t)$ la funzione generalizzata d'appoggio di questa distribuzione:

$$T_D[x] = \int_{-\infty}^{\infty} D(t)x(t)\, dt = -x'(0) \qquad \text{(A.21)}$$

Se calcoliamo la derivata della distribuzione $T_\delta[x]$ abbiamo

$$T'_\delta[x] = -\int_{-\infty}^{\infty} \delta(t)x'(t)\, dt = -x'(0) = T_D[x] \qquad \text{(A.22)}$$

e quindi, ripetendo il ragionamento già fatto a proposito delle funzioni $\delta(t)$ e u(t), concludiamo che la $D(t)$ è la *derivata prima della* $\delta(t)$:

$$D(t) = d\delta(t)/dt \qquad \text{(A.23)}$$

Da questa definizione seguono immediatamente alcune proprietà interessanti della $D(t)$ vista come funzione generalizzata:

$$\int_{-\infty}^{\infty} D(t)x(t-t_0)\, dt = -x'(t_0) \Rightarrow \int_{-\infty}^{\infty} D(t)\, x(t_0 - t)\, dt = x'(t_0) \qquad \text{(A.24)}$$

e quindi

$$x(t) \otimes D(t) = x'(t) \qquad \text{(A.25)}$$

Relazioni simili possono ottenersi definendo le *derivate successive* della funzione $\delta(t)$, e costituiscono una base di analisi dei sistemi lineari e stazionari a un ingresso e una uscita, come sarà chiarito nel Capitolo 4.

A.4 Limite di una distribuzione e di una funzione generalizzata

Consideriamo adesso una *successione* di funzioni generalizzate $\varphi_\varepsilon(t)$. Diremo che questa successione ammette una funzione generalizzata limite $\varphi(t)$, e scriveremo che $\varphi_\varepsilon(t) \underset{\varepsilon\to 0}{\to} \varphi(t)$ quando, per ogni segnale $x(t)$,

$$\lim_{\varepsilon\to 0} T_{\varphi_\varepsilon}[x] = T_\varphi[x] \tag{A.26}$$

ovvero

$$\lim_{\varepsilon\to 0} \int_{-\infty}^{\infty} \varphi_\varepsilon(t)x(t)dt = \int_{-\infty}^{\infty} \varphi(t)x(t)dt \tag{A.27}$$

Questo chiarisce il senso della definizione della funzione generalizzata $\delta(t)$ come limite della successione $\delta_\varepsilon(t)$ (3.4.4). La convergenza di funzioni generalizzate implicata dalla definizione (A.27) non è "diretta" come la convergenza di funzioni ordinarie, in cui si richiede che, per ogni valore di t, si abbia

$$\lim_{\varepsilon\to 0} x_\varepsilon(t) = x(t) \tag{A.28}$$

Nel caso delle funzioni generalizzate, infatti, la convergenza è *mediata* attraverso una operazione di integrazione al di fuori dalla quale la funzione generalizzata stessa *non ha senso*; ricordiamo infatti che quest'ultima ha il solo scopo di definire un *funzionale* (una distribuzione) che a essa si "appoggia". Emerge chiara adesso la necessità di intendere le operazioni al limite che si devono eseguire per calcolare integrali in cui compare la funzione $\delta(t)$ come eseguite *al di fuori* del segno di integrale (si riveda la discussione su questo punto nel Paragrafo 3.4.1). La definizione stessa di convergenza tra funzioni generalizzate (A.27) prevede infatti *esplicitamente* tale operazione.

4

Sistemi monodimensionali a tempo continuo

4.1 Caratterizzazione dei sistemi a tempo continuo

4.1.1 Dal concetto di segnale al concetto di sistema

L'analisi e in generale lo studio dei segnali a tempo continuo intrapresi nei capitoli precedenti originano un certo numero di domande: *In che contesto si manifestano tali segnali? Dove possono essere osservati? Che cosa è responsabile della produzione dei segnali stessi? Come questi possono essere elaborati?* Nei capitoli precedenti abbiamo già visto alcune risposte parziali a queste domande. Abbiamo infatti preso in considerazione segnali prodotti da fenomeni fisici, da circuiti elettrici, da apparati in generale sia naturali che artificiali. Tutti questi esempi possono essere accomunati in un solo concetto, ovvero quello di *sistema.*

Così come nel caso della definizione di segnale discussa nel Capitolo 1, anche la definizione di *sistema* è abbastanza articolata, per lo meno dal punto di vista del linguaggio ordinario. In senso lato, possiamo chiamare *sistema monodimensionale* (altrimenti detto *a un ingresso e una uscita*) un qualunque dispositivo, o interconnessione di dispositivi, o apparato, che produce un segnale di *uscita* (chiamato anche *risposta* o *effetto*) in corrispondenza a un segnale di *ingresso* (detto anche *sollecitazione*, *eccitazione* o *causa*). Questa definizione è intenzionalmente molto vaga, in modo che sotto la dizione *sistema* possano rientrare i casi più disparati. È chiaro che un circuito elettronico per il trattamento del segnale è un caso tipico di sistema (ad esempio un amplificatore in un sistema di riproduzione audio ad alta fedeltà). A buon diritto può però classificarsi come tale anche un sistema di *controllo*: la potenza erogata dal motore a

scoppio di una autovettura (segnale di *uscita*), controllata dal sistema di iniezione di carburante, è determinata dalla posizione che istante per istante assume il pedale dell'acceleratore (segnale di *ingresso*); la temperatura del nocciolo di una centrale termonucleare (segnale di uscita) è determinata dalla portata con cui il liquido refrigerante affluisce al nocciolo stesso (segnale di ingresso), e così via.

Dal punto di vista matematico, che è quello che riguarda più da vicino la teoria dei segnali, la definizione di sistema è assai meno vaga. In questo contesto, un *sistema* è una *trasformazione* (o, con la nomenclatura dell'analisi funzionale, un *funzionale*) che a un segnale di ingresso $x(t)$ fa corrispondere un ben determinato e unico segnale d'uscita $y(t)$. La trasformazione del segnale $x(t)$ nel segnale $y(t)$ si denota nel modo seguente:

$$y(t) = \mathcal{T}[x(\alpha); t] \tag{4.1.1a}$$

ove con questa notazione si intende che il valore dell'uscita all'istante t dipende in generale dall'andamento *complessivo* del segnale d'ingresso $x(t)$, cioè da *tutti* i suoi valori $x(\alpha)$, con $-\infty < \alpha < \infty$ (ad esempio $y(t) = \int_{-\infty}^{t} x(\alpha)\,d\alpha$). Quando non ci sono però particolari questioni di ambiguità, si può usare anche la notazione semplificata

$$y(t) = \mathcal{T}[x(t)] \tag{4.1.1b}$$

Una rappresentazione grafica di questa trasformazione è quella di Figura 4.1. Le frecce indicano i segnali di ingresso e di uscita; il rettangolo è la "materializzazione" grafica della trasformazione $\mathcal{T}[\cdot]$; il punto interrogativo allude al fatto che per il momento del sistema non è nota né la struttura interna (racchiusa dal blocco e inaccessibile), né la maniera per caratterizzarne il comportamento agli effetti esterni. Quest'ultimo argomento è l'oggetto dei prossimi paragrafi.

Figura 4.1 Sistema che trasforma il segnale $x(t)$ nel segnale $y(t)$

Un esempio di sistema è l'*amplificatore ideale* per il quale la legge di trasformazione è elementare: esso viene infatti descritto dalla semplice relazione $y(t) = A\,x(t)$, essendo A una costante data (l'*amplificazione*).

4.1.2 Proprietà dei sistemi monodimensionali

A prescindere dalla struttura interna del sistema, fortemente dipendente dal contesto e dall'applicazione, è possibile acquisire alcune informazioni preliminari sul comportamento del sistema stesso e individuarne così alcune *proprietà,* compiendo osservazioni esclusivamente sui segnali di ingresso/uscita. Esaminiamo dunque queste proprietà in dettaglio.

i) *Stazionarietà*

Se le caratteristiche del sistema *non variano nel tempo*, il sistema è *stazionario*; questo è il caso dei circuiti elettrici con componenti, per l'appunto, costanti nel tempo. Volendo caratterizzare in modo formale un sistema siffatto possiamo scrivere che, se

$$y(t) = \mathcal{T}[x(t)] \tag{4.1.2}$$

allora

$$\mathcal{T}[x(t-t_0)] = y(t-t_0) \tag{4.1.3}$$

Questa relazione dice in pratica che la risposta corrispondente all'eccitazione traslata nel tempo $x(t-t_0)$ ha lo stesso andamento della risposta al segnale originario $x(t)$ non traslato, purché la si trasli della *stessa medesima* quantità t_0.

ii) *Causalità*

Un sistema è *causale* quando il valore dell'uscita all'istante arbitrario generico t dipende soltanto dai valori assunti dall'ingresso agli istanti *precedenti* (o al limite coincidenti con) t stesso:

$$y(t) = \mathcal{T}[x(\alpha), \alpha \le t; t] = \mathcal{T}[x(\alpha)\mathrm{u}(t-\alpha); t] \tag{4.1.4}$$

L'aggettivo *causale* deriva dalla considerazione che, se la relazione precedente non fosse verificata, l'uscita all'istante t sarebbe determinata anche da valori dell'ingresso $x(\alpha)$ a istanti $\alpha > t$, cioè valori *futuri* relativamente a t, in violazione del principio di causa-effetto.

La causalità dei sistemi sembrerebbe quindi una proprietà scontata. In realtà possiamo introdurre un'ulteriore distinzione: si dice che un sistema opera in *tempo reale* se produce il segnale di uscita contestualmente alla presentazione di quello d'ingresso. Quindi un sistema fisicamente realizzabile che lavora in tempo reale non può che essere causale. Se invece l'uscita viene fornita dal sistema solo *successivamente all'acquisizione completa del segnale di ingresso*, si dice che il sistema opera in *tempo virtuale*. Registrando il segnale d'ingresso

su nastro o disco magnetico, si può generare il segnale di uscita elaborando il segnale *successivamente* all'acquisizione (cioè in tempo virtuale) e quindi si possono anche compiere operazioni di tipo "predittivo" impossibili in tempo reale, e tipiche di un sistema non causale.

iii) *Memoria*

Un caso particolare di sistema causale è il cosiddetto sistema *istantaneo* in cui l'uscita all'istante t dipende *solamente* dal valore dell'ingresso al medesimo istante:

$$y(t) = \mathcal{T}[x(\alpha), \alpha = t; t] \tag{4.1.5}$$

In questo caso si usa anche la dizione di sistema *senza memoria* (per evidenti motivi). Due diversi segnali di ingresso $x_1(t)$ e $x_2(t)$, coincidenti a un certo istante t^*, provocano un medesimo valore in uscita all'istante t^*, indipendentemente dai rispettivi andamenti per $t \neq t^*$. L'esempio tipico di sistema istantaneo è l'amplificatore ideale $y(t) = A \cdot x(t)$. Viceversa, un esempio di sistema *con memoria* è il cosiddetto *integratore a finestra mobile* per il quale

$$y(t) = \int_{t-T}^{t} x(\alpha)\, d\alpha \tag{4.1.6}$$

ove $T > 0$ è l'ampiezza della "finestra di integrazione". Il calcolo del valore dell'uscita all'istante t presuppone la conoscenza dell'andamento del segnale d'ingresso in tutto l'intervallo $[t-T, t]$ (Figura 4.2): il sistema mantiene una certa *memoria* dell'andamento del segnale d'ingresso $x(t)$.

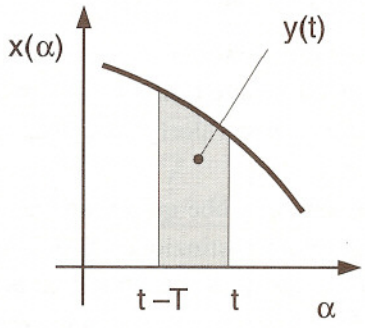


Figura 4.2 Funzionamento di un integratore a finestra mobile

Esempio 4.1

Consideriamo di nuovo l'*amplificatore ideale*. Come già accennato, il suo comportamento è descritto dall'equazione

$$y(t) = A \cdot x(t) \tag{E4.1.1}$$

ove il parametro reale A rappresenta l'amplificazione (cioè il guadagno). Tale sistema è causale e senza memoria. Inoltre esso è stazionario essendo banalmente verificata l'uguaglianza tra $\mathcal{T}\left[x(t-t_0)\right]$ e $y(t-t_0)$.

In una realizzazione dell'amplificatore con componenti elettronici, può accadere che, per lente derive dei componenti stessi, il guadagno non sia costante, ma vari lentamente nel tempo. Un semplice modello di questa situazione è il nuovo sistema

$$y(t) = (A + B\,t)\,x(t) \tag{E4.1.2}$$

che è ancora causale e senza memoria, ma *non è stazionario*. Si ha infatti che

$$\mathcal{T}\left[x(t-t_0)\right] = (A + B\,t)\,x(t-t_0) \neq y(t-t_0) = \left(A + B\,(t-t_0)\right)x(t-t_0). \tag{E4.1.3}$$

❑

iv) *Stabilità*

Diremo che un sistema è *stabile* se, sollecitato da un segnale con andamento arbitrario ma di *ampiezza limitata*, produce a sua volta in uscita un segnale di *ampiezza limitata*:

$$|x(t)| \leq M \Rightarrow |y(t)| \leq K \tag{4.1.7}$$

con M e K finiti. Questa definizione di stabilità si indica con l'acronimo BIBO (*Bounded-Input Bounded-Output*) che significa "uscita limitata per ogni ingresso limitato", ed è solo una tra le molte definizioni di stabilità che si possono dare per i sistemi monodimensionali. Secondo questo criterio, un "buon pilotaggio" di un sistema stabile, cioè un segnale di ingresso adeguatamente limitato, non causa mai in uscita fenomeni di instabilità, cioè situazioni in cui la risposta $y(t)$ tende a crescere illimitatamente.

v) *Invertibilità*

In molti casi è necessario *ricostruire* il segnale di eccitazione in ingresso a un sistema nota la risposta al segnale stesso. Questa operazione è possibile solo per sistemi *invertibili*, per i quali cioè esiste un *sistema inverso* $\mathcal{T}^{-1}[\cdot]$ tale che:

$$\mathcal{T}^{-1}[y(t)] = x(t) \tag{4.1.8}$$

qualunque sia il segnale di ingresso $x(t)$. È chiaro che l'amplificatore ideale è invertibile (e il suo sistema inverso è ancora un amplificatore ideale), mentre il sistema $y(t) = x^2(t)$ non lo è.

vi) *Linearità*
Infine, un sistema è *lineare* se a esso è applicabile il principio di sovrapposizione degli effetti. Ciò significa che al segnale di ingresso

$$x(t) = \alpha x_1(t) + \beta x_2(t) \tag{4.1.9}$$

costituito da una combinazione lineare con coefficienti costanti α e β delle due eccitazioni $x_1(t)$ e $x_2(t)$ (le *cause*), il sistema risponde con il segnale di uscita

$$y(t) = \mathcal{T}[x(t)] = \alpha y_1(t) + \beta y_2(t) \tag{4.1.10}$$

ove $y_1(t) = \mathcal{T}[x_1(t)]$ e $y_2(t) = \mathcal{T}[x_2(t)]$. L'uscita si ottiene dunque mediante la *stessa* combinazione lineare delle due risposte $y_1(t)$ e $y_2(t)$ (gli *effetti*) alle due eccitazioni $x_1(t)$ e $x_2(t)$ considerate agenti separatamente.

Esempio 4.2
Consideriamo il *raddrizzatore a doppia semionda*, la cui relazione costitutiva è

$$\mathcal{T}[x(t)] = |x(t)| \tag{E4.2.1}$$

Poniamo al suo ingresso il segnale $x(t) = x_1(t) + x_2(t)$; poiché

$$\mathcal{T}[x(t)] = |x_1(t) + x_2(t)| \neq y_1(t) + y_2(t) = |x_1(t)| + |x_2(t)| \tag{E4.2.2}$$

si può dire che il sistema in esame non è lineare. ❑

Esempio 4.3
Riprendiamo in considerazione l'amplificatore a guadagno variabile dell'Esempio 4.1:

$$y(t) = (A + Bt)x(t) \tag{E4.3.1}$$

e studiamone la linearità. Se il segnale d'ingresso è $x(t) = \alpha x_1(t) + \beta x_2(t)$, l'uscita corrispondente è data da

$$\mathcal{T}[x(t)] = (A+Bt)[\alpha x_1(t) + \beta x_2(t)] = (A+Bt)\alpha x_1(t) + (A+Bt)\beta x_2(t) \quad \text{(E4.3.2)}$$

Poiché $y_1(t) = (A+Bt)\,x_1(t)$ e $y_2(t) = (A+Bt)\,x_2(t)$ la relazione (E4.3.2) può essere riscritta nella forma

$$y(t) = \alpha y_1(t) + \beta y_2(t) \quad \text{(E4.3.3)}$$

e quindi il sistema è *lineare*. ❑

4.2 Caratterizzazione e analisi dei sistemi lineari stazionari

Restringiamo adesso la nostra attenzione al caso estremamente importante di *sistemi lineari e stazionari* (SLS) (o *invarianti nel tempo)*. Questi rivestono una particolare importanza perché si rivelano estremamente semplici da analizzare, e possono inoltre essere sintetizzati (progettati) con altrettanta facilità.

Nel paragrafo precedente, abbiamo qualificato i sistemi monodimensionali secondo determinate proprietà rilevabili mediante lo studio dei soli segnali di ingresso e uscita, a prescindere dalla struttura materiale del sistema stesso. Vogliamo estendere questo modo di procedere, che potremmo chiamare "a scatola chiusa" con riferimento al diagramma di Figura 4.1, per arrivare a una *caratterizzazione* esaustiva del comportamento dei sistemi *lineari stazionari.*

4.2.1 La risposta impulsiva

Per un SLS dato è possibile misurare (o calcolare se si dispone di uno schema di progetto) la cosiddetta *risposta impulsiva*, cioè l'uscita del sistema in corrispondenza all'eccitazione *impulsiva* $x(t) = \delta(t)$. Convenzionalmente, tale segnale viene indicato con $h(t)$:

$$h(t) \triangleq \mathcal{T}[\delta(t)] \quad (4.2.1)$$

L'importanza della risposta impulsiva di un SLS risiede nel fatto che la sua conoscenza permette di determinare la risposta del sistema a un segnale di ingresso di andamento arbitrario. Ricordiamo la (3.4.21):

$$x(t) = x(t) \otimes \delta(t) = \int_{-\infty}^{\infty} x(\alpha)\,\delta(t-\alpha)\,d\alpha \quad (4.2.2)$$

e scriviamo:

$$y(t) = \mathcal{T}[x(t)] = \mathcal{T}\left[\int_{-\infty}^{\infty} x(\alpha)\delta(t-\alpha)d\alpha\right] \tag{4.2.3}$$

La trasformazione $\mathcal{T}[\cdot]$ caratteristica del sistema e l'operazione di integrale sono entrambe operatori *lineari*, e quindi è possibile invertirne l'ordine di calcolo. La (4.2.3) diventa quindi

$$y(t) = \int_{-\infty}^{\infty} \mathcal{T}[x(\alpha)\delta(t-\alpha)]d\alpha \tag{4.2.4}$$

ove è importante osservare che l'*operatore* $\mathcal{T}[\cdot]$ agisce su segnali *funzioni del tempo t*. Poiché tale operatore è lineare, e tenendo conto che, rispetto al tempo t, la quantità $x(\alpha)$ è una costante, si ha:

$$y(t) = \int_{-\infty}^{\infty} x(\alpha)\cdot\mathcal{T}[\delta(t-\alpha)]d\alpha \tag{4.2.5}$$

e infine, per la proprietà di *stazionarietà* del sistema e ricordando la definizione di risposta impulsiva $h(t)$ (4.2.1), si ottiene:

$$y(t) = \int_{-\infty}^{\infty} x(\alpha)\, h(t-\alpha)\, d\alpha = x(t) \otimes h(t) \tag{4.2.6}$$

che stabilisce la relazione fondamentale (diremmo *costitutiva*) del sistema lineare stazionario: il segnale di uscita può essere calcolato attraverso la *convoluzione* del segnale di ingresso con la risposta impulsiva.

La conoscenza della risposta impulsiva, oltre a permettere di ricavare il segnale di uscita dato quello di ingresso, consente anche di verificare le proprietà possedute dal sistema e quindi *caratterizza completamente* il comportamento del sistema stesso. Nel paragrafo precedente abbiamo visto ad esempio che un sistema è causale se è verificata la relazione

$$y(t) = \mathcal{T}[x(\alpha), \alpha \le t; t] = \mathcal{T}[x(\alpha)\mathrm{u}(t-\alpha); t] \tag{4.2.7}$$

Dimostriamo ora che un SLS è causale se e solo se la sua risposta impulsiva è un *segnale causale* (nel senso specificato nell'Esempio 3.2), cioè

$$h(t) \equiv h(t)\, u(t) \tag{4.2.8}$$

Infatti, se calcoliamo il segnale di uscita di un SLS la cui risposta impulsiva è causale, abbiamo:

$$y(t) = x(t) \otimes h(t) = \int_{-\infty}^{\infty} x(\alpha)\, h(t-\alpha)\, d\alpha = \int_{-\infty}^{\infty} x(\alpha)\, h(t-\alpha) u(t-\alpha)\, d\alpha$$

$$= \int_{-\infty}^{t} x(\alpha)\, h(t-\alpha)\, d\alpha \qquad (4.2.9)$$

in cui l'ultimo passaggio è giustificato dal fatto che $h(t-\alpha) = 0$ per $\alpha \geq t$ come mostrato in Figura 4.3.

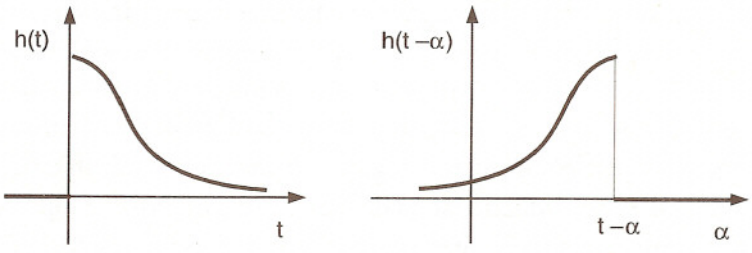


Figura 4.3 Risposta impulsiva di un sistema causale

La limitazione dell'estremo superiore di integrazione nella (4.2.9), provocata dalla causalità della risposta impulsiva, porta a concludere che il valore al generico istante t del segnale di uscita è determinato *dai soli valori assunti da* $x(\alpha)$ *per* $\alpha \leq t$. In una parola, il sistema è causale. Viceversa, se il SLS in esame fosse causale e la (4.2.8) non fosse verificata (cioè la risposta impulsiva *non* fosse causale) si avrebbe un assurdo perché, con un ragionamento analogo a quello appena visto, si dimostrerebbe che l'uscita del sistema a un istante assegnato dipenderebbe anche dai valori assunti dal segnale di ingresso in istanti *successivi*.

Anche la *stabilità* del sistema è univocamente determinata dall'andamento della risposta impulsiva. Infatti, condizione necessaria e sufficiente affinché un SLS sia stabile è che la sua risposta impulsiva sia *assolutamente integrabile*:

$$\int_{-\infty}^{\infty} |h(t)|\, dt < +\infty \qquad (4.2.10)$$

Verifichiamo innanzitutto la sufficienza. Supponiamo dunque

$$\int_{-\infty}^{\infty} |h(t)|\, dt = H < \infty \qquad (4.2.11)$$

Per un segnale d'ingresso limitato (cioè $|x(t)| \le M$) possiamo scrivere

$$|y(t)| = \left| \int_{-\infty}^{\infty} h(\alpha)\, x(t-\alpha)\, d\alpha \right| \le \int_{-\infty}^{\infty} |h(\alpha)\, x(t-\alpha)|\, d\alpha = \int_{-\infty}^{\infty} |h(\alpha)|\, |x(t-\alpha)|\, d\alpha \le$$

$$\le M \int_{-\infty}^{\infty} |h(\alpha)|\, d\alpha = MH < +\infty \qquad (4.2.12)$$

e quindi il sistema è stabile. Procedendo per assurdo dimostriamo poi che la condizione (4.2.10) di assoluta integrabilità è anche necessaria. Supponiamo che $h(t)$ non sia assolutamente integrabile e che il sistema, nonostante ciò, sia stabile. Ciò significa che, dando in ingresso al sistema un segnale limitato arbitrario, il segnale di uscita corrispondente deve avere ampiezza limitata. Questo deve accadere in particolare per il segnale di ingresso (ovviamente limitato) $x(t) = \text{sgn}[h(-t)]$ (il cosiddetto segnale del *caso peggiore*) rappresentato in Figura 4.4 insieme a un particolare andamento della risposta impulsiva $h(t)$.

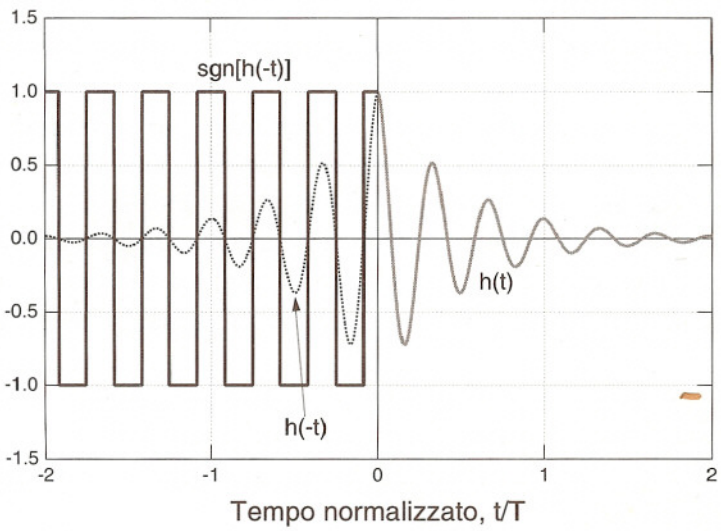


Figura 4.4 Rappresentazione del segnale $x(t) = \text{sgn}[h(-t)]$

Calcoliamo adesso il valore del segnale di uscita $y(t)$ per $t = 0$

$$y(0) = \int_{-\infty}^{\infty} h(\alpha)\, x(t-\alpha)\, d\alpha \Big|_{t=0} = \int_{-\infty}^{\infty} h(\alpha)\, x(-\alpha)\, d\alpha = \int_{-\infty}^{\infty} h(\alpha)\text{sgn}[h(\alpha)]\; d\alpha$$

$$= \int_{-\infty}^{\infty} |h(\alpha)|\, d\alpha = +\infty \tag{4.2.13}$$

avendo supposto, coerentemente con le ipotesi, che la risposta impulsiva del sistema *non* sia assolutamente integrabile. Allora, poiché il segnale di uscita del sistema non ha ampiezza limitata, il sistema in realtà *non è stabile*, in contraddizione con l'ipotesi fatta. Concludendo, possiamo scrivere

$$\text{sistema stabile in senso BIBO} \Leftrightarrow \int_{-\infty}^{\infty} |h(t)|\, dt < +\infty \tag{4.2.14}$$

4.2.2 La risposta in frequenza

La risposta impulsiva di un sistema può essere ricavata applicando in ingresso al sistema stesso un segnale che approssimi la funzione $\delta(t)$ e misurando l'uscita corrispondente. Diciamo "approssimi" perché il segnale $\delta(t)$ è un'astrazione matematica che può solo essere approssimata quando si effettua una misurazione nella pratica. Se però si ha un'idea dei tempi di risposta del sistema, una buona approssimazione della sollecitazione impulsiva è un impulso rettangolare di durata sufficientemente più piccola della costante di tempo intrinseca al sistema, e di ampiezza sufficientemente elevata. Ciò che si ottiene in questo modo è la caratterizzazione del sistema nel *dominio del tempo*.

In molti casi però non è possibile e/o conveniente applicare al sistema una sollecitazione impulsiva, vuoi per l'impossibilità di generare un segnale che sia una buona approssimazione di un impulso di Dirac, vuoi per il timore che una sollecitazione di ampiezza elevata come l'impulso (o meglio, l'approssimazione pratica dell'impulso) possa danneggiare il sistema stesso. Cambiamo dunque tipo di eccitazione, e forniamo al sistema un segnale di ingresso *sinusoidale* o meglio, per semplicità di calcolo, una *oscillazione sinusoidale complessa* alla frequenza f:

$$x(t) = e^{j2\pi ft} \tag{4.2.15}$$

L'uscita corrispondente è espressa da

$$y(t) = \int_{-\infty}^{\infty} h(\alpha)\, x(t-\alpha)\, d\alpha = \int_{-\infty}^{\infty} h(\alpha)\, e^{j2\pi f(t-\alpha)} d\alpha$$

$$= e^{j2\pi ft} \int_{-\infty}^{\infty} h(\alpha)\, e^{-j2\pi f\alpha} d\alpha \tag{4.2.16}$$

Se il sistema è stabile, la risposta a un'oscillazione di frequenza f assegnata è a sua volta un'oscillazione alla stessa frequenza f, ma *modificata in ampiezza e fase* rispetto all'ingresso di un fattore a valori complessi che chiamiamo *risposta in frequenza* (talvolta *risposta armonica*) del sistema:

$$H(f) \triangleq \left. \frac{y(t)}{x(t)} \right|_{x(t)=e^{j2\pi ft}} \tag{4.2.17}$$

Al variare della frequenza f, la variazione di ampiezza e lo sfasamento introdotto dal sistema su un segnale sinusoidale cambiano, cosicché la risposta in frequenza deve intendersi come una *funzione della frequenza f* esprimibile ad esempio in modulo e fase. Questa definizione di carattere *operativo* della risposta in frequenza può essere presentata in modo alternativo se si riconsidera la (4.2.16):

$$H(f) = \int_{-\infty}^{\infty} h(\alpha)\, e^{-j2\pi f\alpha} d\alpha \tag{4.2.18}$$

Allora è chiaro che la risposta in frequenza è anche ricavabile come *trasformata di Fourier della risposta impulsiva*:

$$H(f) \triangleq \mathcal{F}[h(t)] \tag{4.2.19}$$

Le due definizioni (4.2.17) e (4.2.19) sono chiaramente equivalenti. La prima però suggerisce un metodo per *misurare* la risposta in frequenza attraverso segnali di prova sinusoidali a frequenza variabile. La seconda richiede la preventiva conoscenza della risposta impulsiva, e quindi ha come prerequisito la caratterizzazione *temporale* del sistema, che non è richiesta viceversa nella "misura" della risposta in frequenza (4.2.17).

Se infine indichiamo con $X(f)$ e $Y(f)$ le trasformate di Fourier rispettivamente del segnale di ingresso e di quello di uscita, possiamo trovare una terza maniera di definire o ricavare la risposta in frequenza $H(f)$. Trasformando entrambi i membri della relazione costitutiva del SLS (4.2.6) abbiamo

$$y(t) = x(t) \otimes h(t) \Leftrightarrow Y(f) = X(f)\, H(f) \tag{4.2.20}$$

ovvero

$$H(f) = \frac{Y(f)}{X(f)} \tag{4.2.21}$$

Questa relazione permette di ricavare la risposta in frequenza del SLS attraverso la conoscenza delle trasformate di Fourier dell'ingresso e dell'uscita del sistema, prescindendo ancora una volta da una misura della risposta impulsiva. È chiaro però che quest'ultima definizione ha senso solo per quei valori di f per cui $X(f) \neq 0$. Le tre definizioni viste per $H(f)$ possono essere usate indifferentemente nella risoluzione di problemi pratici o teorici a seconda della convenienza.

Sappiamo che la trasformata di Fourier di un segnale reale gode della proprietà di simmetria Hermitiana (si vedano le (3.2.8-3.2.9)); perciò se la risposta impulsiva di un sistema è reale possiamo scrivere:

$$H(f) = H^*(-f) \Rightarrow \begin{cases} A(f) = A(-f) \\ \theta(f) = -\theta(-f) \end{cases} \tag{4.2.22}$$

dove $A(f) = |H(f)|$ è la cosiddetta *risposta in ampiezza* del sistema e $\theta(f) = \angle H(f)$ è la sua *risposta in fase*. Per comprendere appieno il perché di questa nomenclatura (analoga peraltro a quella adottata per gli spettri dei segnali), poniamo in ingresso al sistema il segnale

$$x(t) = a\cos(2\pi f_0 t + \varphi_0) \tag{4.2.23}$$

e calcoliamo il corrispondente segnale d'uscita $y(t)$. Osserviamo che $x(t)$ può essere riscritto come

$$x(t) = \frac{a}{2} e^{j\varphi_0} e^{j2\pi f_0 t} + \frac{a}{2} e^{-j\varphi_0} e^{-j2\pi f_0 t} \tag{4.2.24}$$

Pertanto, sfruttando la definizione di risposta in frequenza (4.2.17) e la linearità del sistema si ottiene:

$$\begin{aligned} y(t) &= H(f_0)\frac{a}{2} e^{j\varphi_0} e^{j2\pi f_0 t} + H(-f_0)\frac{a}{2} e^{-j\varphi_0} e^{-j2\pi f_0 t} \\ &= A(f_0)\, e^{j\theta(f_0)} \frac{a}{2} e^{j\varphi_0} e^{j2\pi f_0 t} + A(-f_0)\, e^{j\theta(-f_0)} \frac{a}{2} e^{-j\varphi_0} e^{-j2\pi f_0 t} \\ &= A(f_0)\frac{a}{2}\left[e^{j(\theta(f_0)+\varphi_0)} e^{j2\pi f_0 t} + e^{-j(\theta(f_0)+\varphi_0)} e^{-j2\pi f_0 t}\right] \\ &= a\, A(f_0)\cos(2\pi f_0 t + \varphi_0 + \theta(f_0)) \end{aligned} \tag{4.2.25}$$

Come già ricavato nel caso dell'oscillazione complessa, l'uscita del sistema risulta modificata in ampiezza e fase relativamente al segnale d'ingresso in ragione rispettivamente delle *risposte in ampiezza e fase* del sistema alla frequenza f_0 considerata.

Esempio 4.4

Consideriamo la consueta squadra *R-C* mostrata in Figura 4.5 e determiniamo la sua risposta in frequenza $H(f)$. Osserviamo preliminarmente che il problema è ben posto, nel senso che il sistema dato è lineare e stazionario poiché costituito da componenti lineari a valori costanti nel tempo. La risposta impulsiva del sistema può essere ricavata come segue:

$$h(t) = \mathcal{T}[\delta(t)] = \mathcal{T}\left[\frac{d\,\mathrm{u}(t)}{dt}\right] = \frac{d}{dt}\mathcal{T}[\mathrm{u}(t)] = \frac{dg(t)}{dt} \qquad \text{(E4.4.1)}$$

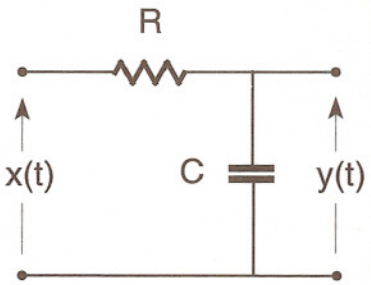


Figura 4.5 Circuito elettrico *R-C*

ove $g(t)$ è come di consueto la risposta del sistema al gradino unitario $u(t)$. Ancora una volta nella (E4.4.1) sono state invertite le operazioni *lineari* di derivazione temporale e trasformazione operata dal sistema.

La risposta al gradino del circuito può essere interpretata come il risultato di un'operazione di "carica del condensatore *C*" attraverso la resistenza *R* (si veda a questo proposito la Figura 3.29). Nozioni di fisica elementare permettono di scrivere

$$g(t) = \left(1 - e^{-\frac{t}{RC}}\right)\mathrm{u}(t) = \left(1 - e^{-t/T}\right)\mathrm{u}(t) \qquad \text{(E4.4.2)}$$

relazione rappresentata in Figura 4.6, e dove $T = RC$ è la costante di tempo del circuito. Derivando, si ottiene la risposta impulsiva $h(t)$ rappresentata in Figura 4.7:

$$h(t) = \frac{dg(t)}{dt} = \frac{1}{T}e^{-t/T}u(t) \qquad \text{(E.4.4.3)}$$

La sua trasformata di Fourier è

$$H(f) = \frac{1}{1 + j2\pi f T} \tag{E4.4.4}$$

che rappresenta infine la risposta in frequenza cercata del sistema dato.

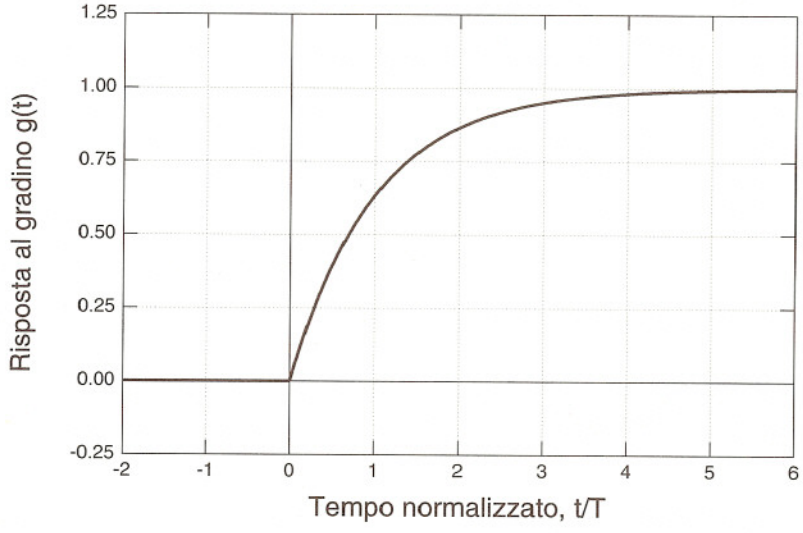


Figura 4.6 Risposta al gradino del circuito *R-C*

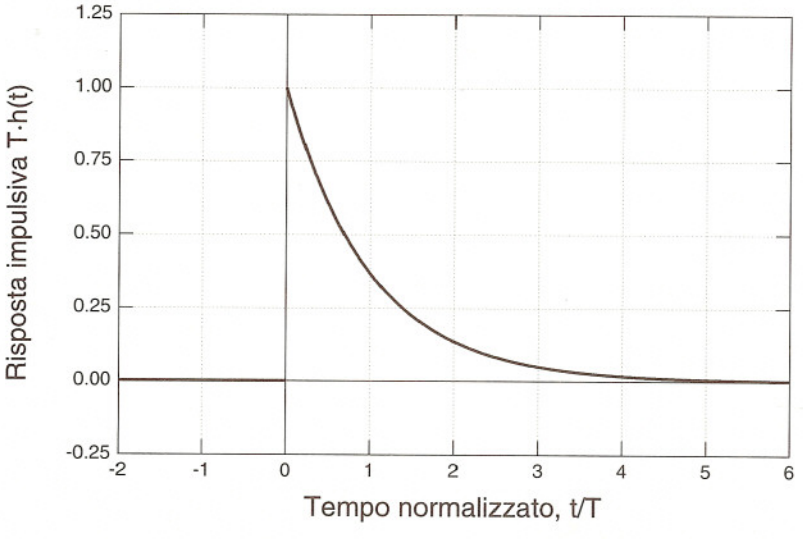


Figura 4.7 Risposta impulsiva del circuito *R-C*

❑

Esempio 4.5

Determiniamo la risposta in frequenza del circuito di Figura 4.5 per altra via. A tal fine facciamo riferimento alla notazione indicata in Figura 4.8.

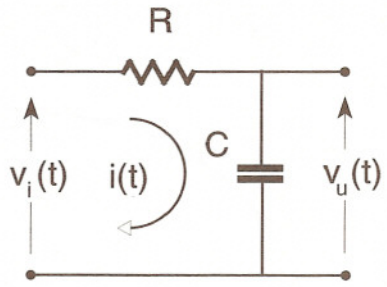


Figura 4.8 Circuito elettrico *R-C*

Considerando la caduta di tensione sul resistore e la tensione ai capi del condensatore si ha

$$v_u(t) = v_i(t) - R\, i(t) \tag{E4.5.1}$$

Indicando con $q(t)$ la carica elettrica accumulata sulle armature del condensatore fino all'istante t, si ha inoltre

$$v_u(t) = \frac{q(t)}{C} \quad \Rightarrow \frac{d\, v_u(t)}{dt} = \frac{1}{C}\frac{dq(t)}{dt} = \frac{i(t)}{C} \tag{E4.5.2}$$

per cui

$$i(t) = C\frac{d\, v_u(t)}{dt} \tag{E4.5.3}$$

e, sostituendo questa relazione nella (E4.5.1), si ottiene

$$v_u(t) = v_i(t) - RC\frac{d\, v_u(t)}{dt} \tag{E4.5.4}$$

che descrive la relazione tra i segnali di ingresso $x(t) = v_i(t)$ e di uscita $y(t) = v_u(t)$. Calcolando la trasformata di Fourier di entrambi i membri della (E4.5.4) si ha

$$V_u(f) = V_i(f) - j2\pi f RCV_u(f) \tag{E4.5.5}$$

da cui si ricava

$$H(f)=\frac{V_u(f)}{V_i(f)}=\frac{1}{1+j2\pi f RC} \qquad \text{(E4.5.6)}$$

Se si definisce la *frequenza di taglio* del sistema (il cui significato sarà chiarito più avanti) come

$$f_T=\frac{1}{2\pi RC} \qquad \text{(E4.5.7)}$$

la risposta in frequenza del SLS è anche

$$H(f)=\frac{1}{1+j\,f/f_T} \qquad \text{(E4.5.8)}$$

dalla quale si trovano immediatamente le seguenti espressioni della risposta in ampiezza e in fase:

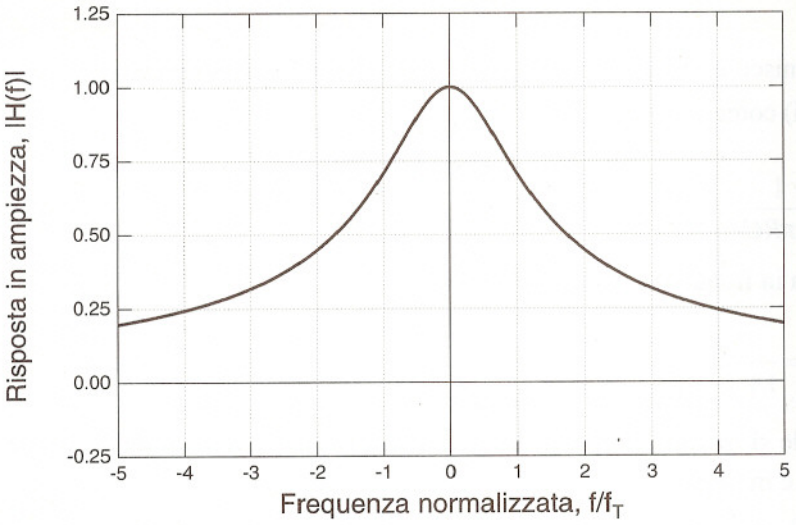
$$|H(f)|=\frac{1}{\sqrt{1+(f/f_T)^2}} \quad , \quad \angle H(f)=-\text{arctg}\frac{f}{f_T} \qquad \text{(E4.5.9)}$$

L'andamento delle risposte in ampiezza e fase è rappresentato nella Figura 4.9 mediante scale *lineari* per entrambi gli assi del riferimento cartesiano del grafico. ❑

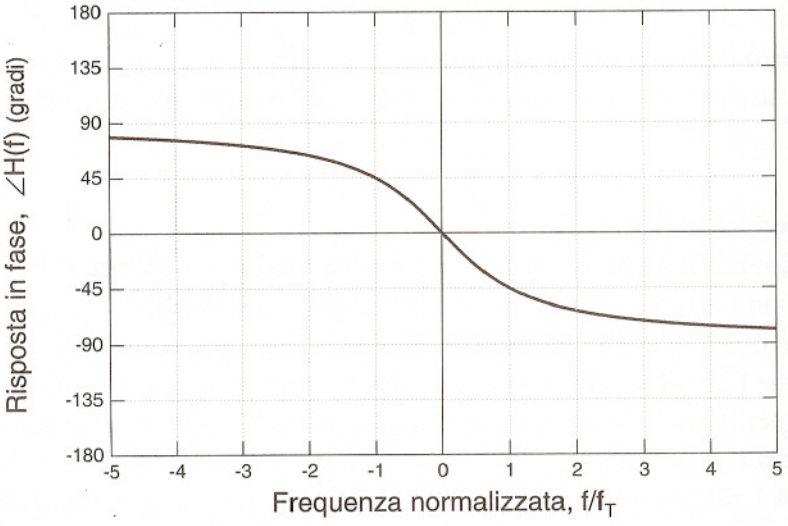
4.2.3 Il decibel

In molti fenomeni fisici, specialmente collegati con la percezione sensoriale umana, l'ampiezza di un segnale può variare di molti ordini di grandezza. L'esempio tipico di questa tendenza è dato dai segnali acustici udibili. I suoni più forti che l'orecchio umano può tollerare senza danni hanno una intensità di *pressione acustica I* pari a circa 1 W/m^2 (ad esempio, un motore aeronautico a piena potenza). I più deboli rumori percepibili, viceversa, hanno una pressione acustica di 1 pW/m^2= 10^{-12} W/m^2, cioè 12 ordini di grandezza inferiore. Quel che più conta inoltre è che a *sensazione* di intensità acustica raddoppiata, triplicata ecc. rispetto a una intensità di riferimento I_0 corrisponde in realtà una *effettiva* intensità acustica che varia come $(I/I_0)^2$, $(I/I_0)^3$ ecc. In questi casi dunque è appropriato usare una scala di misura *logaritmica,* in cui al "raddoppio soggettivo" della intensità corrisponde un raddoppio dell'unità di misura, che

deve essere intesa come logaritmo (in base 10) del fattore di incremento reale. Questa unità è il *bel* (abbreviato B) o più, comunemente, il *decibel* (abbreviato dB) pari evidentemente a un decimo di bel.



(a)



(b)

Figura 4.9 Risposta in ampiezza (a) e fase (b) del circuito R-C

Per quel che riguarda la misura della risposta in ampiezza di un sistema, si definisce la risposta *in dB* del sistema come

$$|H(f)|_{dB} \triangleq 10 \log_{10} \frac{|H(f)|^2}{|H(f_0)|^2} \tag{4.2.26}$$

Poiché il dB, come chiarito poc'anzi, è sempre e comunque una misura *relativa a un riferimento*, nella definizione compare il fattore $|H(f_0)|^2$ che costituisce appunto il valore rispetto al quale i dB sono calcolati; la frequenza f_0 del riferimento deve essere scelta opportunamente per i vari tipi di sistema considerati, come sarà chiarito in seguito; il moltiplicatore 10 serve infine a convertire i bel (dati dal solo logaritmo del rapporto) in decibel. La risposta in ampiezza del sistema viene elevata al quadrato perché il dB si riferisce a misure relative di *potenza*, e la potenza dei segnali trattati dal SLS è legata alla risposta in ampiezza *elevata al quadrato* (si veda il Paragrafo 4.4).

Esempio 4.6

Ricaviamo la risposta in ampiezza del SLS dell'Esempio 4.5 espressa in decibel:

$$|H(f)|_{dB} = 10 \log_{10} \frac{|H(f)|^2}{|H(f_0)|^2} = 10 \log_{10} |H(f)|^2 = -10 \log_{10} \left[1 + \left(\frac{f}{f_T} \right)^2 \right] \tag{E4.6.1}$$

Abbiamo scelto come riferimento la frequenza $f_0 = 0$ per la quale si ha il valore massimo della risposta in ampiezza: $|H(f_0)| = 1$. Spesso è utile una scala logaritmica *anche per l'asse delle frequenze* espresse in Hz. Ciò permette di visualizzare l'andamento della risposta in un ambito molto esteso senza perdere il dettaglio sulle basse frequenze. In tal caso viene rappresentata la sola parte della risposta per $f > 0$, e si ottiene il grafico di Figura 4.10a. In Figura 4.10b è rappresentata inoltre la risposta in fase del sistema ancora su scala frequenziale logaritmica. ❑

Esempio 4.7

Determiniamo la risposta in frequenza del circuito *C-R* di Figura 4.11. Da un esame dello schema si ha

$$v_u(t) = R\, i(t) \quad , \quad v_u(t) = v_i(t) - \frac{q(t)}{C} \tag{E4.7.1}$$

ove $q(t)$ rappresenta la carica nel condensatore all'istante t. Combinando le due equazioni e derivando rispetto al tempo si ricava

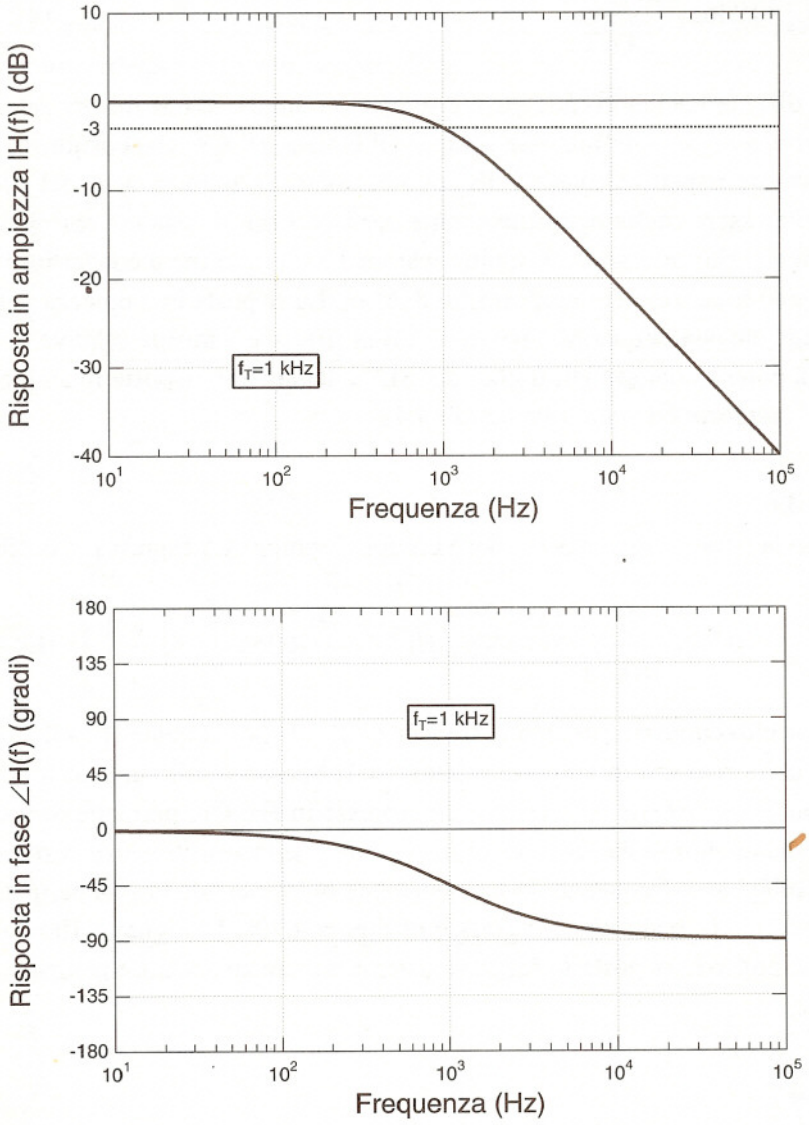


Figura 4.10 Risposta in ampiezza in dB (a) e in fase (b) del circuito *R-C*

$$\frac{d}{dt}v_i(t) - \frac{i(t)}{C} = R\frac{d}{dt}i(t) \qquad \text{(E4.7.2)}$$

Poiché inoltre $i(t) = v_u(t)/R$ si ottiene l'equazione differenziale che lega i segnali di ingresso e di uscita:

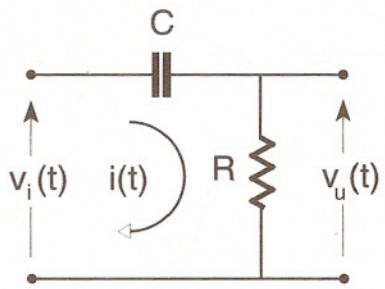


Figura 4.11 Circuito elettrico *C-R* (Esempio 4.7)

$$\frac{d}{dt}v_i(t)-\frac{v_u(t)}{RC}=\frac{d}{dt}v_u(t) \qquad \text{(E4.7.3)}$$

Calcolando la trasformata di Fourier di entrambi i membri si ha

$$j2\pi f\ V_i(f)-\frac{1}{RC}V_u(f)=j2\pi f V_u(f) \qquad \text{(E4.7.4)}$$

da cui si ricava la risposta in frequenza del sistema:

$$H(f)=\frac{V_u(f)}{V_i(f)}=\frac{j2\pi f RC}{1+j2\pi f RC}=\frac{j2\pi f T}{1+j2\pi f T} \quad , \quad T=RC \qquad \text{(E4.7.5)}$$

Definendo anche in questo caso la *frequenza di taglio* del circuito

$$f_T=\frac{1}{2\pi RC} \qquad \text{(E4.7.6)}$$

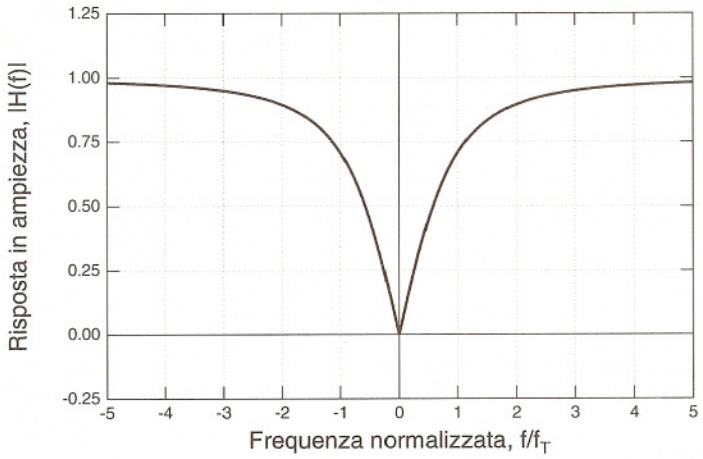
la risposta in frequenza del sistema diventa

$$H(f)=\frac{j\left(f/f_T\right)}{1+j\left(f/f_T\right)}=1-\frac{1}{1+j\left(f/f_T\right)} \qquad \text{(E4.7.7)}$$

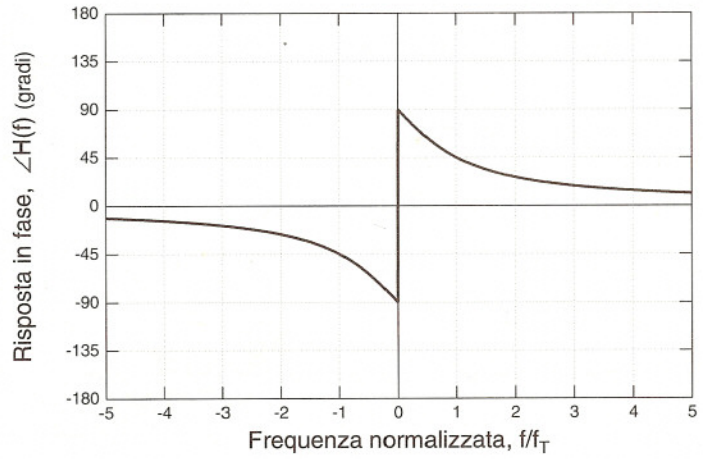
cui corrispondono le risposte in ampiezza e fase

$$|H(f)|=\sqrt{\frac{\left(f/f_T\right)^2}{1+\left(f/f_T\right)^2}} \quad , \quad \angle H(f)=\frac{\pi}{2}\text{sgn(f)}-\text{arctg}\frac{f}{f_T}=-\text{arctg}\frac{f_T}{f} \qquad \text{(E4.7.8)}$$

che sono rappresentate in Figura 4.12 utilizzando scale lineari per entrambi gli assi di riferimento.



(a)



(b)

Figura 4.12 Risposta in ampiezza (a) e fase (b) del circuito *C-R*

Prima di calcolare l'espressione della risposta in ampiezza del circuito in dB, osserviamo che in questo caso non ha senso scegliere come riferimento $f_0 = 0$, in quanto $H(0) = 0$. Viceversa, visto l'andamento della risposta in ampiezza, si deve scegliere stavolta $f_0 \to \infty$, per cui

$$|H(f)|_{dB} = 10 \log_{10} \frac{|H(f)|^2}{|H(\infty)|^2} = 10 \log_{10} \left[\frac{(f/f_T)^2}{1+(f/f_T)^2} \right] \qquad \text{(E4.7.9)}$$

L'andamento di $|H(f)|_{dB}$ è illustrato in Figura 4.13 insieme al grafico della risposta in fase, entrambi su scala logaritmica per le frequenze.

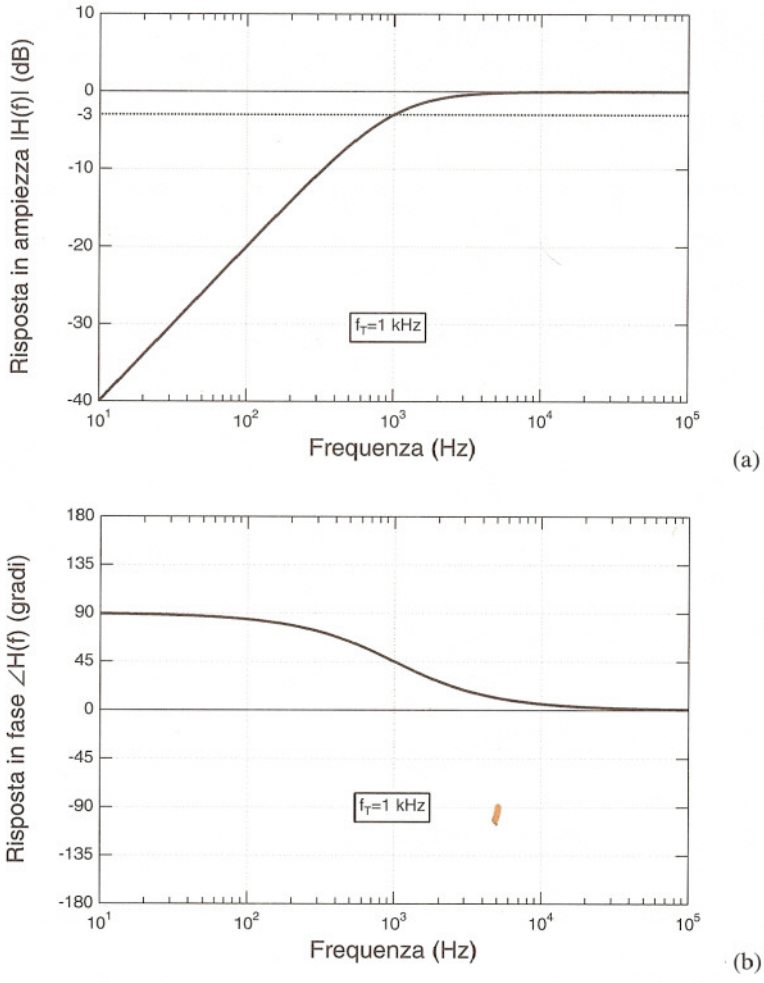


Figura 4.13 Risposta in ampiezza misurata in dB (a) e in fase (b) del circuito *C-R*

Antitrasformando la risposta in frequenza (E4.7.7) si ricava la risposta impulsiva del circuito:

$$h(t) = \delta(t) - \frac{1}{T} e^{-t/T} \, \mathrm{u}(t) \qquad \text{(E4.7.10)}$$

che è rappresentata in Figura 4.14 ($T = RC$). Dalla risposta impulsiva si può ricavare infine la risposta al gradino $g(t)$:

$$g(t) = \int_{-\infty}^{t} h(\alpha)\, d\alpha = \int_{-\infty}^{t} \left[\delta(\alpha) - \frac{1}{T} e^{-t/T} \, \mathrm{u}(\alpha) \right] d\alpha = \mathrm{u}(t) - \frac{1}{T} \int_{-\infty}^{t} e^{-\alpha/T} \, \mathrm{u}(\alpha)\, d\alpha$$

$$= \mathrm{u}(t) - \frac{\mathrm{u}(t)}{T}\int_0^t e^{-\alpha/T}\, d\alpha = e^{-t/T}\,\mathrm{u}(t) \qquad \text{(E4.7.11)}$$

rappresentata a sua volta in Figura 4.15.

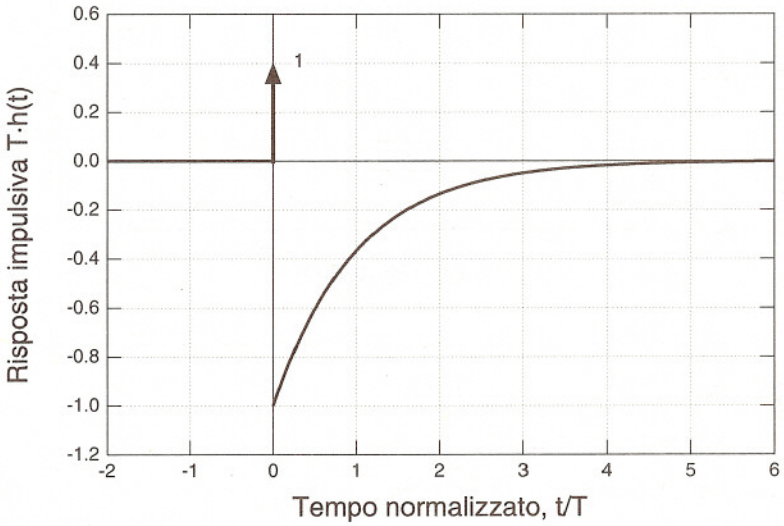


Figura 4.14 Risposta impulsiva del circuito *C-R*

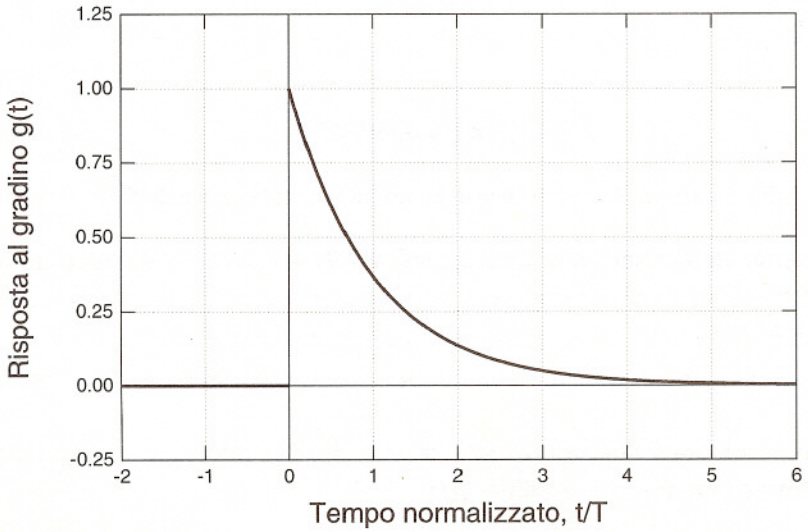


Figura 4.15 Risposta al gradino del circuito *C-R*

❑

Esempio 4.8

Determiniamo la risposta $y(t)$ del circuito di Figura 4.16 al cui ingresso è applicato il segnale $x(t)$ di Figura 4.17.

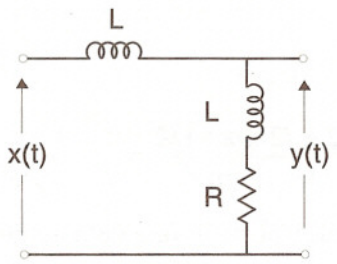


Figura 4.16 Circuito elettrico dell'Esempio 4.8

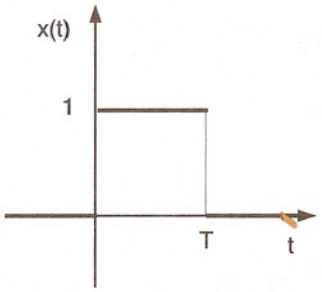


Figura 4.17 Segnale di ingresso del circuito di Figura 4.16

Si osservi che il segnale in ingresso può essere espresso nella forma

$$x(t) = \mathrm{u}(t) - \mathrm{u}(t-T) \tag{E4.8.1}$$

da cui segue immediatamente

$$y(t) = g(t) - g(t-T) \tag{E4.8.2}$$

dove $g(t)$ è come di consueto la risposta al gradino unitario del circuito in esame. Per determinare $g(t)$ ricaviamo la risposta impulsiva $h(t)$ del sistema come *antitrasformata* della risposta in frequenza $H(f)$. Consideriamo dunque il segnale di eccitazione $x(t) = e^{j2\pi ft}$; attraverso il calcolo fasoriale è semplice trovare l'espressione del corrispondente segnale di uscita $y(t)$:

$$y(t) = e^{j2\pi ft}\frac{z_L + R}{2z_L + R} = e^{j2\pi ft}\frac{j2\pi fL + R}{2j2\pi fL + R} \tag{E4.8.3}$$

ove evidentemente z_L è l'impedenza dell'induttore L alla frequenza f. È immediato ricavare l'espressione della risposta in frequenza cercata:

$$H(f)=\frac{R+j2\pi fL}{R+j4\pi fL} \tag{E4.8.4}$$

da cui, ponendo $\alpha = L/R$, si ricava

$$H(f)=\frac{1+j2\pi f\alpha}{1+j4\pi f\alpha}=\frac{1/2+j2\pi f\alpha+1/2}{1+j4\pi f\alpha}=\frac{1}{2}+\frac{1}{2}\frac{1}{1+j4\pi f\alpha} \tag{E4.8.5}$$

La risposta in ampiezza e la risposta in fase del circuito sono rappresentate nella Figura 4.18. Antitrasformando la (E4.8.5) si ottiene l'espressione della risposta impulsiva del circuito:

$$h(t)=\frac{1}{2}\delta(t)+\frac{1}{4\alpha}e^{-t/2\alpha}u(t) \tag{E4.8.6}$$

e quindi la risposta al gradino è

$$g(t)=\mathrm{u}(t)\otimes h(t)=\int_{-\infty}^{t}h(\tau)\,d\tau=\frac{1}{2}\,\mathrm{u}(t)+\frac{1}{4\alpha}\int_{-\infty}^{t}e^{-\tau/2\alpha}\,\mathrm{u}(\tau)\,d\tau$$

$$=\frac{1}{2}\,\mathrm{u}(t)+\mathrm{u}(t)\frac{1}{4\alpha}\int_{0}^{t}e^{-\tau/2\alpha}\,d\tau=\left[1-\frac{1}{2}e^{-t/2\alpha}\right]\mathrm{u}(t) \tag{E4.8.7}$$

per cui il segnale di uscita risulta

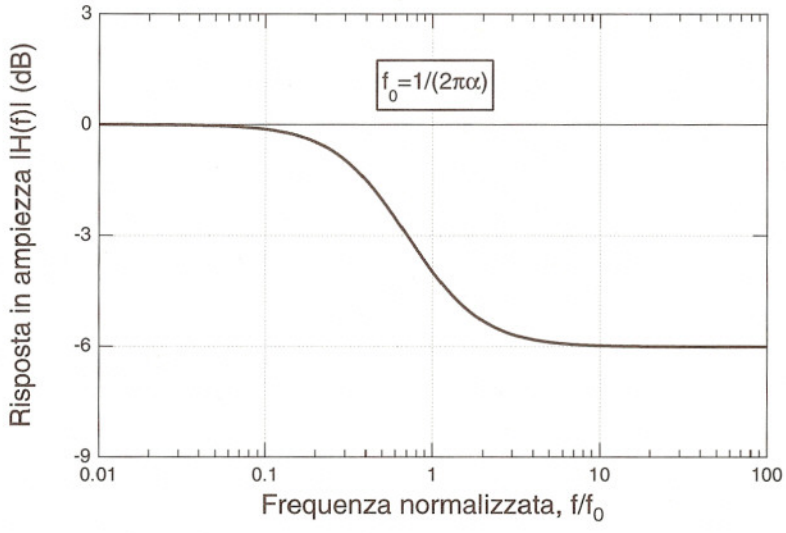
$$y(t)=\left[1-\frac{1}{2}e^{-t/2\alpha}\right]\mathrm{u}(t)-\left[1-\frac{1}{2}e^{-(t-T)/2\alpha}\right]\mathrm{u}(t-T) \tag{E4.8.8}$$

e cioè

$$y(t)=\begin{cases}0 & t\le 0\\ 1-\dfrac{1}{2}e^{-t/2\alpha} & 0<t\le T\\ \dfrac{1}{2}\left(e^{T/2\alpha}-1\right)e^{-t/2\alpha} & t>T\end{cases} \tag{E4.8.9}$$

segnale che è visualizzato in Figura 4.19. È interessante notare che, se $T \ll \alpha$, allora:

$$y(t)\approx\frac{1}{2}\,\mathrm{rect}\left(\frac{t-T/2}{T}\right)=\frac{1}{2}\,x(t) \tag{E4.8.10}$$



(a)

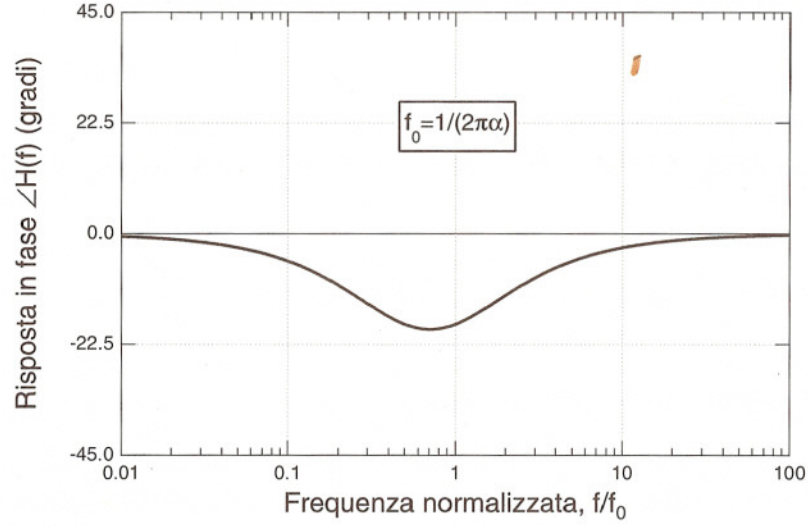


Figura 4.18 Risposta in ampiezza (a) e in fase (b) del sistema dell'Esempio 4.8

mentre, se $T >> \alpha$,

$$y(t) \approx \text{rect}\left(\frac{t - T/2}{T}\right) = x(t) \qquad \text{(E4.8.11)}$$

L'andamento di entrambi i due casi limite (E4.8.10-E4.8.11) si può facilmente riscontrare rispettivamente dalle curve per $\alpha = 2T$ e $\alpha = T/8$ di Figura 4.19.

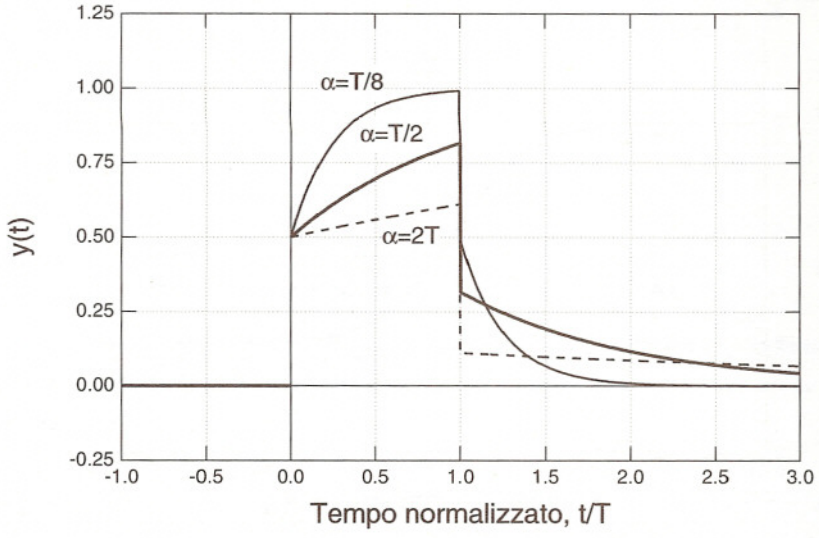


Figura 4.19 Risposta del sistema di Figura 4.16 al segnale di Figura 4.17

❑

4.2.4 Sistemi in cascata e in parallelo

Consideriamo ora due SLS stabili disposti *in cascata*, come illustrato in Figura 4.20.

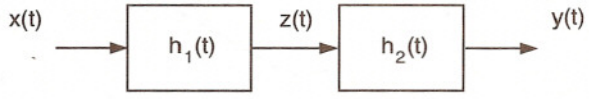


Figura 4.20 Sistemi lineari stazionari in cascata

Indicando con $h_1(t)$ e $h_2(t)$ le risposte impulsive rispettivamente del primo e del secondo sistema, vogliamo determinare la relazione esistente tra il segnale $x(t)$ in ingresso al primo sistema e il segnale $y(t)$ in uscita al secondo. Sappiamo che

$$y(t) = \int_{-\infty}^{\infty} z(\alpha)\, h_2(t-\alpha)\, d\alpha \tag{4.2.27}$$

e che

$$z(t) = \int_{-\infty}^{\infty} x(\beta)\, h_1(t-\beta)\, d\beta \tag{4.2.28}$$

Sostituendo la (4.2.28) nella (4.2.27) si ottiene:

$$y(t) = \int_{\alpha=-\infty}^{\infty} h_2(t-\alpha)\left[\int_{\beta=-\infty}^{\infty} h_1(\alpha-\beta)\,x(\beta)\,d\beta\right]d\alpha$$

$$= \int_{\beta=-\infty}^{\infty} x(\beta)\left[\int_{\alpha=-\infty}^{\infty} h_1(\alpha-\beta)\,h_2(t-\alpha)\,d\alpha\right]d\beta \qquad (4.2.29)$$

e con l'ulteriore sostituzione $\gamma = \alpha - \beta$ nell'integrale in $d\alpha$,

$$y(t) = \int_{\beta=-\infty}^{\infty} x(\beta)\left[\int_{\gamma=-\infty}^{\infty} h_1(\gamma)\,h_2\big((t-\beta)-\gamma\big)\,d\gamma\right]d\beta$$

$$= \int_{-\infty}^{\infty} x(\beta)\big[h_1(t-\beta)\otimes h_2(t-\beta)\big]\,d\beta = x(t)\otimes\big[h_1(t)\otimes h_2(t)\big] \qquad (4.2.30)$$

Questo risultato dimostra che la cascata di Figura 4.20 può essere rappresentata come un *unico sistema equivalente* con risposta impulsiva

$$h(t) = h_1(t)\otimes h_2(t) \qquad (4.2.31)$$

ovvero risposta in frequenza

$$H(f) = H_1(f)\,H_2(f) \qquad (4.2.32)$$

come suggerito dalla Figura 4.21. Nel procedimento analitico che porta al risultato (4.2.31) è stata fatta la tacita ipotesi che i due sistemi in cascata *non si influenzino a vicenda*, cioè che il comportamento dei due (in particolare la loro risposta impulsiva), quando vengono connessi in cascata, sia identico a quello riscontrato per ciascuno isolatamente. Questa ipotesi *non* è verificata ad esempio per i circuiti *R-C* e *C-R* esaminati negli Esempi 4.4-4.5. In tal caso è necessario interporre eventualmente tra i due un *circuito disaccoppiatore* (o *buffer*) che impedisce reciproche influenze tra i due stadi.

Un secondo tipo molto comune di interconnessione è quella in *parallelo*, in cui i sistemi vengono alimentati dallo stesso ingresso, e le uscite dei due vengono poi sommate, come in Figura 4.22. È chiaro che la risposta impulsiva e in frequenza equivalenti sono in questo caso pari a

$$h(t) = h_1(t) + h_2(t) \quad , \quad H(f) = H_1(f) + H_2(f) \qquad (4.2.33)$$

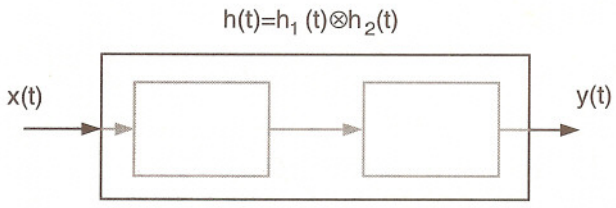


Figura 4.21 Sistema equivalente alla cascata di Figura 4.20

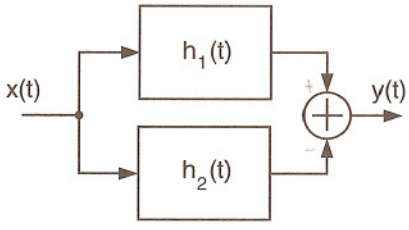


Figura 4.22 Sistemi in parallelo

4.3 Filtri

4.3.1 Generalità sui filtri e filtri ideali

Un caso tipico che si presenta nell'elaborazione dei segnali è quello in cui il segnale osservato $x(t)$ è costituito dalla sovrapposizione, cioè dalla *somma*, di due segnali: $x(t) = x_1(t) + x_2(t)$ dei quali il primo è un segnale utile, cioè portatore di informazione, mentre il secondo rappresenta solo un *disturbo* ineliminabile alla fonte. Nella raccolta di un segnale elettrocardiografico, ad esempio, può accadere che alla tensione raccolta dai sensori sul corpo del paziente (molto debole, dell'ordine di grandezza dei mV e stilizzata in Figura 4.23a) venga a sovrapporsi un disturbo dovuto alla *tensione di alimentazione* fornita all'apparato elettromedicale dalla normale rete elettrica 220 V-50 Hz (come in Figura 4.23b). Se il circuito elettrico dello strumento non è realizzato con la massima accuratezza, il residuo della tensione di alimentazione può rivelarsi dello stesso ordine di grandezza del segnale utile. In un caso di questo genere, è fondamentale riuscire a discriminare il segnale utile dal disturbo, cosa apparentemente impossibile tenendo conto che il segnale osservato è la *sovrapposizione* di queste due componenti, come si vede dalla Figura 4.24.

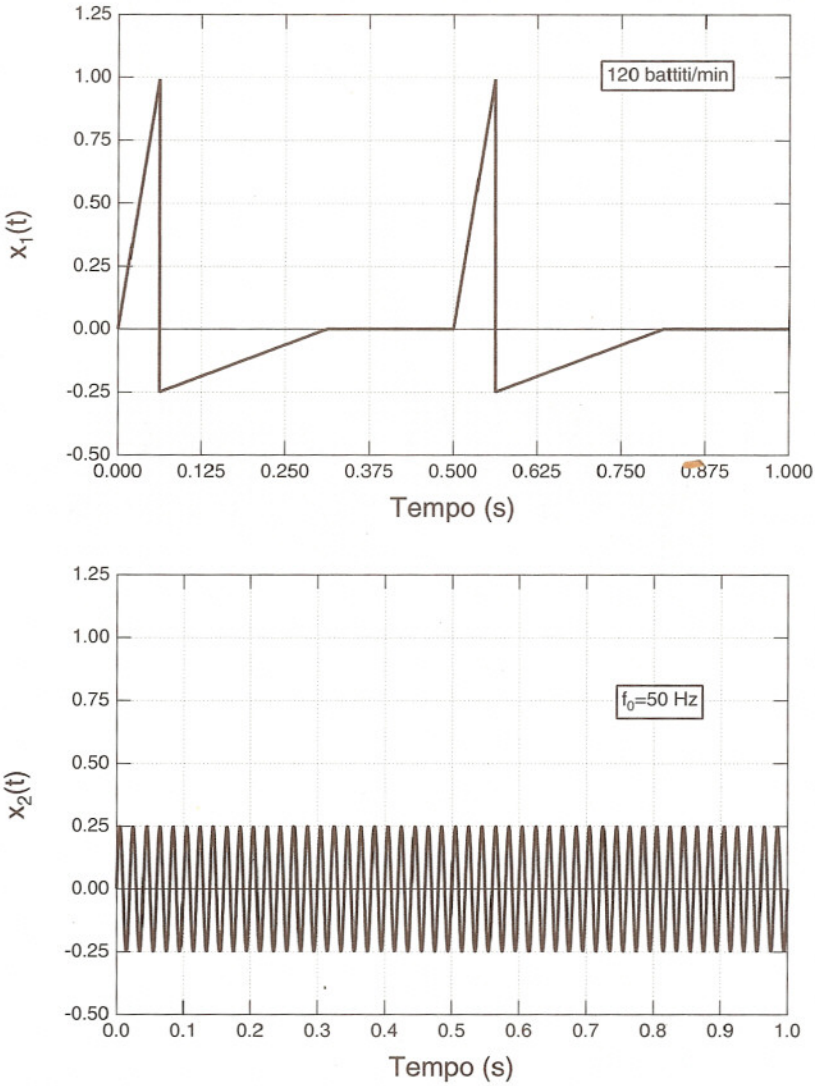


Figura 4.23 Segnale utile (a) e disturbo (b) in un elettrocardiogramma

Consideriamo però i segnali in *ambito frequenziale*. La situazione è quella rappresentata in Figura 4.25, in cui lo spettro del segnale "utile" e quello del "disturbo" *insistono su intervalli frequenziali disgiunti.* Si intuisce allora che è possibile s*eparare il segnale utile dal disturbo* utilizzando un SLS con risposta in frequenza opportuna. Se, come di consueto, indichiamo con $X_1(f)$ e $X_2(f)$ le trasformate di Fourier rispettivamente di $x_1(t)$ e $x_2(t)$, la trasformata di Fourier

del segnale $x(t)$ è allora

$$X(f) = X_1(f) + X_2(f) \tag{4.3.1}$$

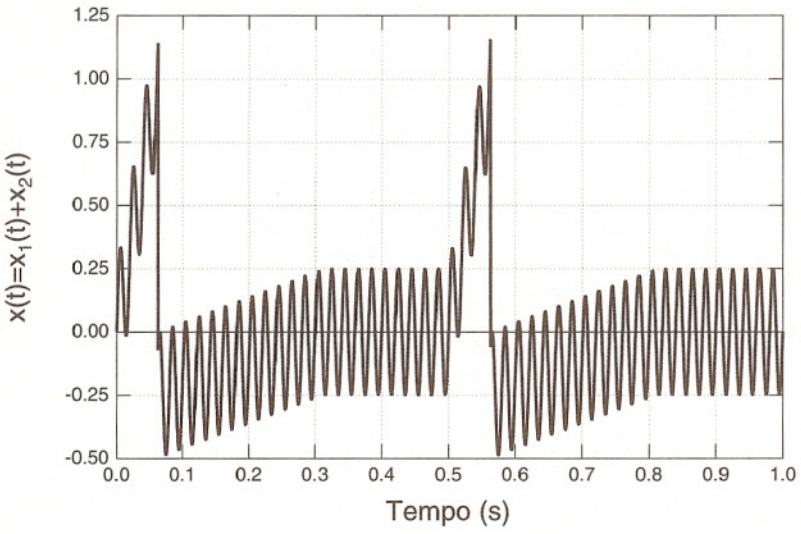


Figura 4.24 Sovrapposizione del segnale e del disturbo di Figura 4.23

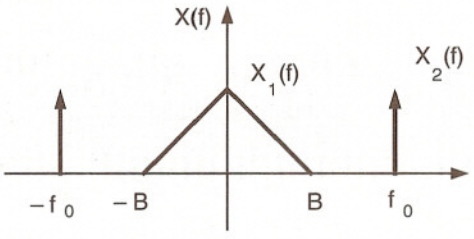


Figura 4.25 Spettro del segnale $x(t)$ di Figura 4.24

Se vogliamo *reiettare* (cioè cancellare) il disturbo $x_2(t)$, possiamo elaborare il segnale tramite un SLS con caratteristiche di *selettività* nei confronti delle varie componenti frequenziali che compongono il segnale. In particolare, è evidente che il segnale viene preservato e il disturbo viene reiettato se il sistema ha una risposta in frequenza pari a

$$H_{LP}(f) = \text{rect}\left(\frac{f}{2B}\right) \tag{4.3.2}$$

illustrata in Figura 4.26. Il segnale d'uscita $y(t)$ del filtro, avendo infatti trasformata di Fourier $Y(f) = X(f)\,H(f)$, sarà privo del disturbo $x_2(t)$.

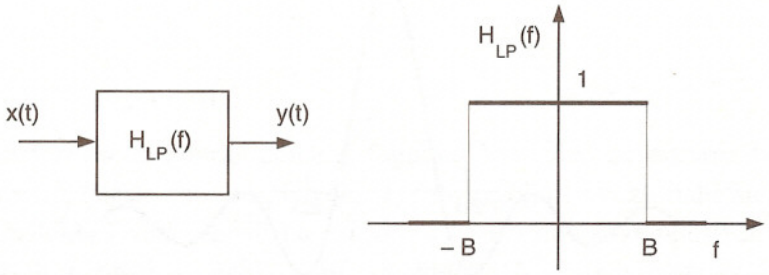


Figura 4.26 Risposta in frequenza di un *filtro passa-basso ideale*

Un sistema con risposta in frequenza come in Figura 4.26 viene chiamato *filtro passa-basso ideale*. Esso infatti possiede caratteristiche di selettività, nel senso che le componenti frequenziali all'interno di una certa *banda*, cioè intervallo di frequenze, vicino alla frequenza nulla (quindi *basse* frequenze) vengono lasciate inalterate. In questa zona infatti, chiamata *banda passante*, si ha $H(f) = 1$. Viceversa, all'esterno della banda passante, e cioè nella cosiddetta *banda oscura*, le componenti frequenziali vengono *completamente cancellate* perché $H(f) = 0$. La frequenza B rappresenta il cosiddetto *limite di banda*.

Questa funzione di selettività giustifica il nome di "filtro" dato a questo SLS, nel senso che le componenti nello spettro del segnale aventi frequenza maggiore del limite di banda vengono "trattenute", mentre le altre componenti vengono "lasciate passare". Nella pratica, si tende ad identificare il limite di banda B con l'ampiezza della banda passante (o banda *tout-court*). Per convenzione, infatti, la *banda* del filtro è *l'ampiezza* della banda passante *considerata sul solo semiasse positivo delle frequenze*. Per il filtro di Figura 4.26, quindi, la banda è pari a B e coincide con il limite di banda.

La risposta impulsiva del filtro passa-basso (*low-pass*) ideale si ricava antitrasformando l'espressione della risposta in frequenza (4.3.2):

$$h_{LP}(t) = 2B\,\mathrm{sinc}(2Bt) \tag{4.3.3}$$

funzione rappresentata in Figura 4.27. Si nota immediatamente che $h_{LP}(t)$ è diversa da zero anche per valori di $t < 0$, per cui il filtro passa-basso ideale è un sistema *non causale* e quindi fisicamente non realizzabile.

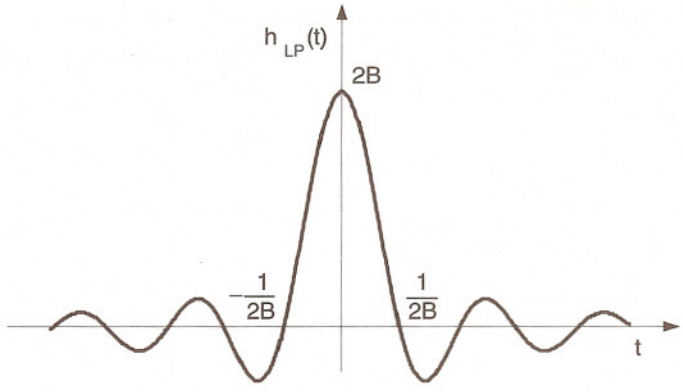


Figura 4.27 Risposta impulsiva del filtro passa-basso ideale

Se invece desiderassimo sopprimere la componente $x_1(t)$ nel segnale $x(t)$ per isolare $x_2(t)$, potremmo utilizzare un SLS la cui risposta in frequenza è rappresentata in Figura 4.28. Tale sistema, che permette l'eliminazione delle basse frequenze, viene chiamato *filtro passa-alto ideale* ed è caratterizzato dalle risposte in frequenza e impulsiva seguenti:

$$H_{HP}(f) = 1 - \text{rect}\left(\frac{f}{2B}\right) \Leftrightarrow h_{HP}(t) = \delta(t) - 2B\text{sinc}(2Bt) \tag{4.3.4}$$

È chiaro che stavolta la banda passante del filtro passa alto (*high-pass*) è quella che sta *al di là* del limite di banda B, nella quale le componenti frequenziali del segnale di ingresso non vengono alterate. Colloquialmente, diremo ancora (in modo improprio) che la banda del filtro passa-alto è B, alludendo in realtà al *limite* di banda. È immediato verificare che anche il filtro passa-alto ideale è un sistema non causale.

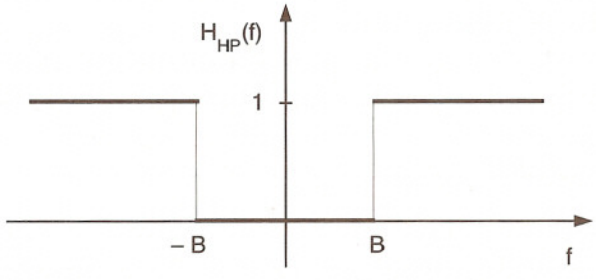


Figura 4.28 Risposta in frequenza di un filtro passa alto ideale

Nella situazione di Figura 4.25 la separazione tra i due segnali $x_1(t)$ e $x_2(t)$ è stata effettuata rispettivamente con un filtro passa-basso o con un filtro passa-alto ideali. Supponiamo invece di osservare il segnale $x(t)$ somma di *tre* componenti

$$x(t) = x_1(t) + x_2(t) + x_3(t) \tag{4.3.5}$$

aventi trasformate di Fourier come in Figura 4.29, e di voler estrarre da esso il segnale $x_2(t)$; questo caso è tipico dei segnali *modulati* emessi dalle stazioni di radiodiffusione. I vari segnali hanno infatti spettri non sovrapposti in ambito frequenziale e posti a cavallo delle cosiddette *frequenze portanti* su cui i radioricevitori vengono poi *sintonizzati*. Come mostra la figura, è necessario disporre di un sistema con risposta in frequenza $H_{BP}(f)$ non nulla solo nella banda occupata dal segnale $x_2(t)$. Tale sistema prende il nome di *filtro passa-banda ideale*. La banda passante del filtro (ripetiamo, definita per convenzione sul solo semiasse delle frequenze positive) si estende tra il limite di banda *inferiore* f_L e il limite di banda *superiore* f_H.

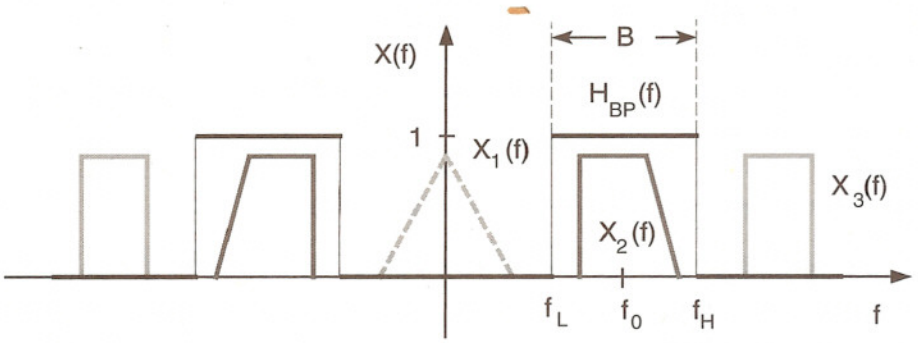


Figura 4.29 Spettro del segnale e risposta in frequenza del filtro passa-banda ideale

Alternativamente ai limiti di banda, il filtro passa-banda (*band-pass*) viene più comunemente caratterizzato attraverso i due parametri equivalenti *frequenza centrale* (o *di centro-banda*) $f_0 = (f_L + f_H)/2$, e ampiezza della banda passante $B = f_H - f_L$. Calcoliamo ora la risposta impulsiva del filtro passa-banda ideale ricordando il *teorema della modulazione*:

$$x(t)\cos(2\pi f_0 t) \Leftrightarrow \frac{X(f - f_0) + X(f + f_0)}{2} \tag{4.3.6}$$

Nel nostro caso,

$$H_{BP}(f) = \text{rect}\left(\frac{f - f_0}{B}\right) + \text{rect}\left(\frac{f + f_0}{B}\right) \tag{4.3.7}$$

e quindi

$$h_{BP}(t) = 2B\,\text{sinc}(Bt)\cos(2\pi f_0 t) \tag{4.3.8}$$

relazione illustrata nella Figura 4.30 nel caso in cui $f_0 = 10B$.

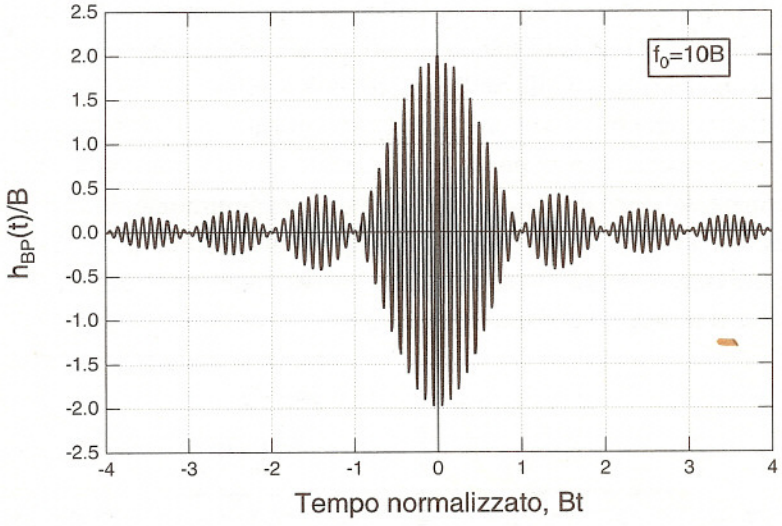


Figura 4.30 Risposta impulsiva di un filtro passa-banda ideale

Ricordiamo infine che per i filtri passa-banda si definisce anche un altro parametro, detto *fattore di qualità Q*, che mette in relazione la frequenza centrale con la banda del filtro:

$$Q \triangleq \frac{f_0}{B} \tag{4.3.9}$$

e che è tanto maggiore quanto minore è la banda passante relativamente alla frequenza centrale, cioè quanto più il filtro è *selettivo*.

Consideriamo di nuovo il segnale $x(t)$ (4.3.5) il cui spettro è rappresentato in Figura 4.31, e supponiamo di voler *eliminare* da esso il segnale $x_2(t)$. Il filtro che permette di effettuare tale operazione viene detto *filtro elimina-banda ideale* e la sua risposta in frequenza $H_{BR}(f)$ è rappresentata in Figura 4.31. Tale risposta è ancora caratterizzata da una frequenza centrale f_0 e da un'ampiezza di banda B (o dai limiti di banda f_L ed f_H), entrambe però relative alla banda

oscura. È immediato stabilire la seguente relazione tra la risposta in frequenza del filtro elimina-banda (*band-reject*) ideale e quella del filtro passa banda ideale:

$$H_{BR}(f) = 1 - H_{BP}(f) \tag{4.3.10}$$

per cui

$$h_{BR}(t) = \delta(t) - 2B\,\mathrm{sinc}(Bt)\cos(2\pi f_0 t) \tag{4.3.11}$$

Un filtro elimina-banda può essere molto selettivo, e al limite può essere utilizzato per reiettare la componente spettrale a una *sola* particolare frequenza (ad esempio, la frequenza di rete $f_0 = 50$ Hz). In tal caso il filtro elimina-banda viene più comunemente chiamato filtro *notch*.

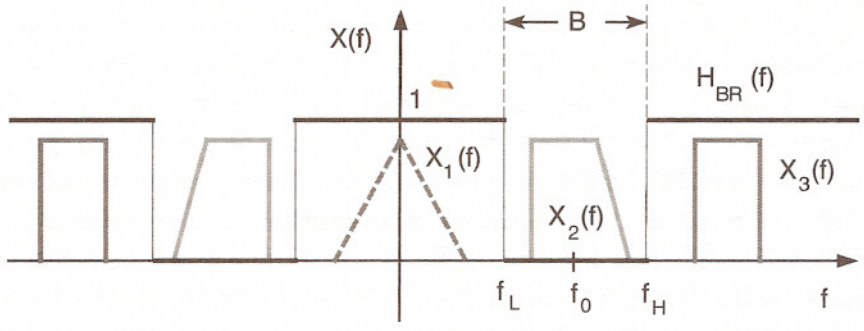


Figura 4.31 Risposta in frequenza di un filtro elimina-banda ideale

4.3.2 Criterio di Paley-Wiener e filtri reali

Tutti i filtri ideali che sono stati appena presi in esame sono *non causali* perché hanno risposte impulsive non nulle per $t < 0$. Questa non-causalità emerge chiara da un esame delle caratteristiche *temporali* dei sistemi considerati. Tuttavia, anche nota la sola risposta *in frequenza* di un SLS, è possibile decidere se essa è relativa a un sistema causale o meno. A questo proposito è utile il *criterio di Paley-Wiener*, che riguarda i sistemi lineari stazionari la cui risposta in ampiezza è a quadrato integrabile:

$$\int_{-\infty}^{\infty} |H(f)|^2 df < \infty \tag{4.3.12}$$

e che quindi è applicabile direttamente ai filtri passa-basso e passa-banda ideali.

In questa ipotesi, se la risposta in ampiezza $|H(f)|$ è tale da verificare la ulteriore condizione

$$\int_{-\infty}^{\infty} \frac{|\ln[|H(f)|]|}{1+(2\pi f)^2} df < \infty \tag{4.3.13}$$

allora esiste una funzione reale $\theta(f)$ tale che

$$H(f) = |H(f)| e^{j\theta(f)} \tag{4.3.14}$$

rappresenta la risposta in frequenza di un sistema causale; se viceversa la condizione (4.3.13) non è verificata, il sistema *non è causale*. Questa condizione è quindi *necessaria* alla causalità del sistema, ma non è sufficiente. In particolare, il criterio non indica come scegliere l'opportuna $\theta(f)$ che rende effettivamente il sistema causale.

Nel caso dei filtri ideali passa-basso e passa-banda, che verificano l'ipotesi di applicabilità del criterio, notiamo che le risposte in ampiezza sono *nulle su intervalli frequenziali di misura (ampiezza) non nulla.* Questo implica che l'integrale della (4.3.13) è sicuramente *divergente* perché la funzione integranda è non limitata su intervalli di misura non nulla. Il criterio non è quindi soddisfatto, e i filtri sono necessariamente non causali. Indirettamente, questo dice che anche i filtri passa-alto ed elimina-banda sono non causali in quanto le loro risposte in frequenza sono il complemento a uno delle risposte rispettivamente del passa-basso e passa-banda aventi gli stessi limiti di banda.

Conseguenza di queste osservazioni è che un filtro causale può avere una risposta in frequenza che si annulla *solo per un insieme di frequenze di misura nulla,* cioè soltanto per un numero finito (o infinito numerabile) di frequenze. Esempi elementari di filtri passa-basso e passa-alto *reali* sono rispettivamente le squadre *R-C* e *C-R* di Figura 4.8 e 4.11 per le quali le risposte in ampiezza sono

$$|H_{R-C}(f)| = \frac{1}{\sqrt{1+(f/f_T)^2}} \quad , \quad |H_{C-R}(f)| = \frac{|f/f_T|}{\sqrt{1+(f/f_T)^2}} \tag{4.3.15}$$

la prima delle quali resta sempre $\neq 0$, mentre la seconda si annulla nel solo punto $f = 0$, senza violare il criterio di Paley-Wiener.

Filtri elementari di tipo passa-banda ed elimina-banda si possono realizzare con circuiti risonanti *L-C* rispettivamente come in Figura 4.32a-b, ove $2\pi f_0 = 1/\sqrt{LC}$ e $1/Q = 2\pi f_0 L/R$. Le risposte in frequenza di questi filtri sono:

$$H_{L-C/BP}(f) = \frac{jQ^{-1} f/f_0}{1-(f/f_0)^2 + jQ^{-1} f/f_0} \tag{4.3.16a}$$

$$H_{L-C/BR}(f) = \frac{1-(f/f_0)^2}{1-(f/f_0)^2 + jQ^{-1} f/f_0} \tag{4.3.16b}$$

le cui ampiezze sono rappresentate in Figura 4.33.

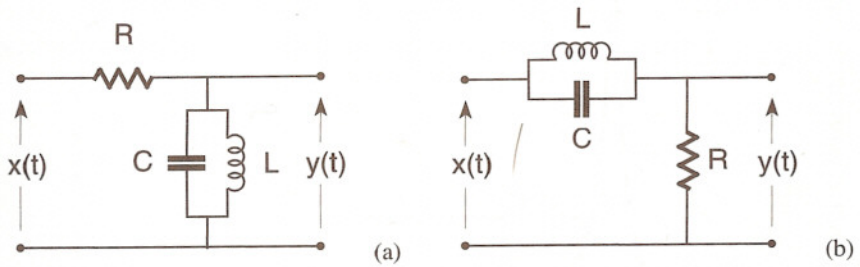


Figura 4.32 Filtri passa-banda (a) ed elimina-banda (b) reali con gruppi *L-C* parallelo

4.3.3 Banda e durata di un segnale e banda di un sistema

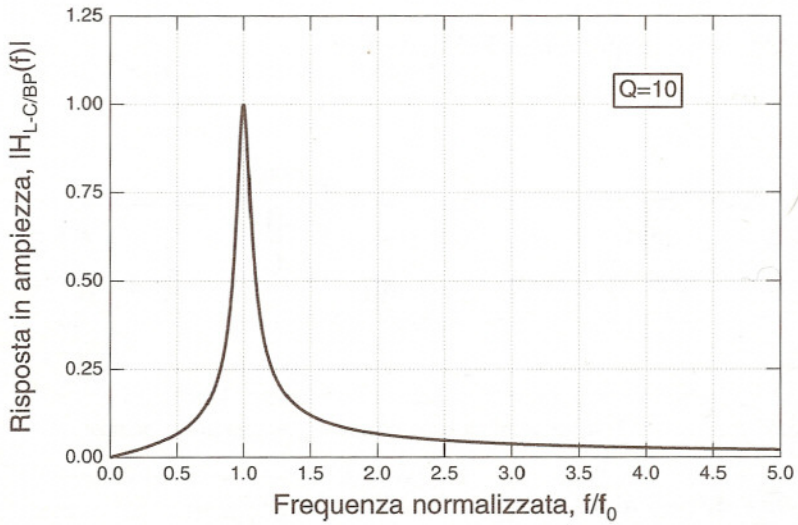
Le caratteristiche selettive dei filtri "reali" sono chiaramente meno spiccate di quelle dei filtri ideali, nel senso che per essi non è individuabile esattamente una banda passante e una oscura come per i prototipi ideali. Altrettanto chiaro però è che essi possiedono tuttavia un certo grado di selettività che ne rende utile l'impiego in pratica come sistemi filtranti. In questi casi è necessario definire *convenzionalmente* i limiti di banda dei filtri reali, e all'atto pratico considerare questi limiti convenzionali come del tutto analoghi a quelli propriamente detti relativi ai filtri ideali (cioè non fisicamente realizzabili).

Prima di arrivare alle definizioni convenzionali di banda di un sistema, premettiamo una discussione sui concetti generali di *durata* e *banda* di un segnale, che sono stati prima d'ora ampiamente utilizzati nel testo senza una definizione rigorosa. Consideriamo dapprima un segnale $x(t)$ *a durata finita*, cioè non nullo solo su di un intervallo limitato $[t_1, t_2]$ (Figura 4.34). La *durata* del segnale è in questo caso pari a $D = t_2 - t_1$. Per il segnale a durata finita esiste un certo valore T (precisamente $T = \max(|t_1|, |t_2|)$) tale che il segnale originario non subisce alcuna modifica se viene moltiplicato per una funzione rect(·) di durata $2T$:

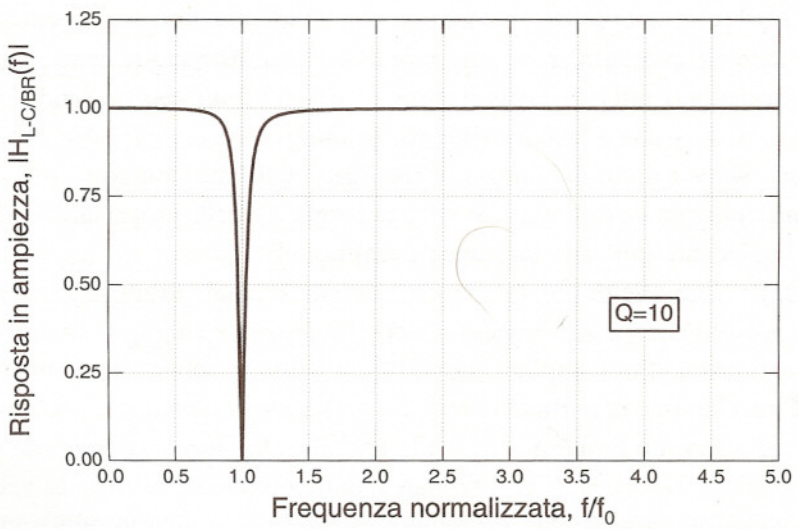
$$x(t) = x(t)\,\mathrm{rect}\left(\frac{t}{2T}\right) \tag{4.3.17}$$

Passando alle trasformate di Fourier, si ha

$$X(f) = X(f) \otimes 2T \operatorname{sinc}(2fT) \tag{4.3.18}$$



(a)



(b)

Figura 4.33 Risposte in ampiezza dei filtri *L-C* passa-banda (a) ed elimina-banda (b)

Questa relazione permette di rispondere (anche se in modo non del tutto rigoroso) alla seguente domanda: può il segnale $x(t)$ a durata finita avere anche *banda* finita, cioè avere trasformata di Fourier diversa da zero solo su di un

intervallo di frequenze limitato? Se rispondiamo di sì a questa domanda cadiamo in un assurdo perché il secondo membro della (4.3.18) è la convoluzione tra due trasformate delle quali la seconda è una funzione sinc(·) avente *estensione* (cioè *banda*) *illimitata.* Il risultato della convoluzione ha quindi estensione (banda) illimitata, in contrasto con la nostra affermazione.

Dunque *un segnale a durata rigorosamente limitata ha banda infinita.* Dualmente, *un segnale con banda rigorosamente limitata ha durata illimitata.* Molti segnali non hanno però né banda né durata rigorosamente limitata, cosicché è spesso necessario definire *convenzionalmente* queste due quantità.

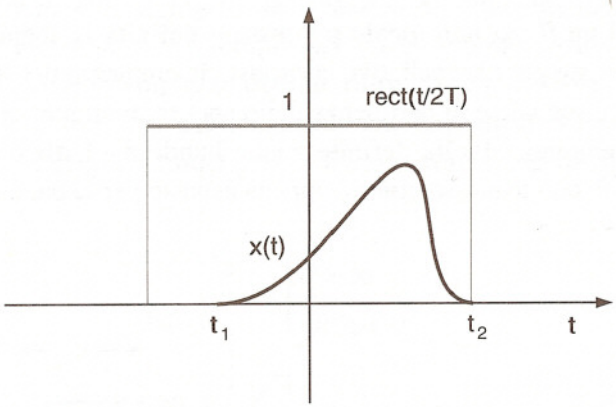


Figura 4.34 Segnale a durata finita

Consideriamo un segnale $x(t)$ il cui spettro di ampiezza assume il valore massimo per $f = f_0$. Se lo spettro è di tipo passa-basso, $f_0 = 0$; se lo spettro è passa-alto, $f_0 = +\infty$, altrimenti lo spettro è di tipo passa-banda e f_0 è finita e diversa da zero. Per gli spettri passa-basso o passa-alto, si definisce *limite di banda a* -3 *dB* quella frequenza per la quale lo spettro di ampiezza risulta ridotto di un fattore $\sqrt{2}$ rispetto al valore di riferimento (Figura 4.35a[1]):

$$\frac{|X(B_{-3})|}{|X(f_0)|} = \frac{1}{\sqrt{2}} \Rightarrow \frac{|X(B_{-3})|^2}{|X(f_0)|^2} = \frac{1}{2} \tag{4.3.19}$$

ovvero

[1] Nella figura vengono mostrati gli spettri di ampiezza dei segnali solo per $f > 0$ perché la definizione di banda riguarda sempre il solo semiasse positivo delle frequenze.

$$10 \log_{10} \frac{|X(B_{-3})|^2}{|X(f_0)|^2} = 10 \log_{10} \frac{1}{2} = -10 \log_{10} 2 \approx -3 \text{ dB} \qquad (4.3.20)$$

risultato che giustifica il nome di *banda a –3 dB*. Quest'ultima rappresenta una indicazione "pratica" della banda di un segnale, nel senso che, per uno spettro passa-basso, le componenti a frequenze inferiori a B_{-3} sono arbitrariamente ritenute "importanti" nella composizione del segnale, mentre quelle oltre B_{-3} sono ritenute "trascurabili". Per gli spettri passa-banda, si possono identificare *due* limiti di banda a –3 dB e una *banda* a –3 dB come differenza tra i due limiti (Figura 4.35b).

La banda di un *filtro* non ideale può essere definita in maniera analoga a quanto appena visto per i segnali, ove la risposta in ampiezza del sistema gioca il ruolo che per un segnale è di pertinenza dello spettro di ampiezza. In particolari applicazioni vengono talvolta definite anche bande a –1 dB o a –20 dB dei segnali e sistemi, con lo stesso criterio appena esposto per la banda a –3 dB.

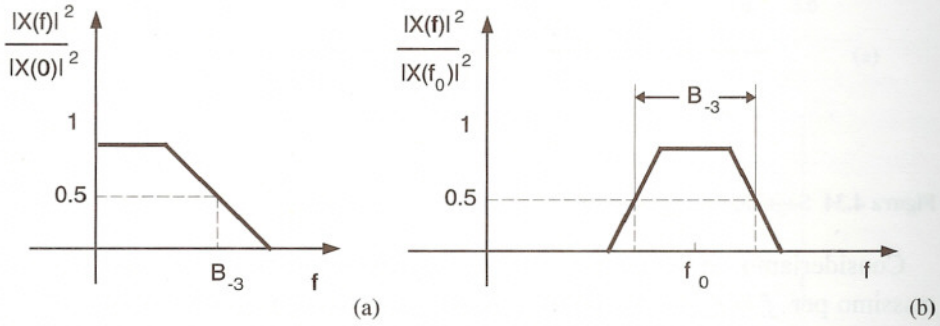


Figura 4.35 Banda a –3 dB di uno spettro passa-basso (a) e passa-banda (b)

La definizione più comune di *durata* convenzionale di un segnale è molto simile a quella di banda per un segnale passa-banda. Se t_0 rappresenta l'istante in cui il segnale presenta il valore di riferimento (generalmente quello di massimo modulo, come in Figura 4.36), si identificano gli istanti t_1 e t_2 $(t_2 > t_1)$ in cui il segnale assume in valore assoluto la *metà* del valore di riferimento $|x(t_0)|$ e si definisce come *durata a metà ampiezza* del segnale la quantità $D = t_2 - t_1$.

Esempio 4.9

Riprendiamo in considerazione il filtro passa-basso reale implementato come squadra *R-C* degli Esempi 4.4-4.5. La sua risposta in frequenza è

$$H(f)=\frac{1}{1+j2\pi f\,RC} \tag{E4.9.1}$$

Calcoliamone la banda a –3 dB. Essendo il filtro passa-basso, scegliamo come valore della frequenza di riferimento $f_0=0$, in corrispondenza della quale si ha $H(0)=1$. Quindi la banda B_{-3} è data da

$$|H(B_{-3})|^2=1/2 \Rightarrow \frac{1}{1+(2\pi B_{-3}RC)^2}=1/2 \Rightarrow B_{-3}=\frac{1}{2\pi RC} \tag{E4.9.2}$$

cioè la banda a –3 dB coincide con quella che negli Esempi 4.4-4.5 era stata chiamata la *frequenza di taglio* del sistema. Si dimostra facilmente che la squadra passa-alto *C-R* dell'Esempio 4.7 ha questo stesso limite di banda. I grafici delle risposte in ampiezza dei due filtri *R-C* e *C-R* rispettivamente nelle Figure 4.10 e 4.13 riportano esplicitamente una linea di riferimento a –3 dB che mette in evidenza geometricamente il risultato analitico appena ricavato. ❑

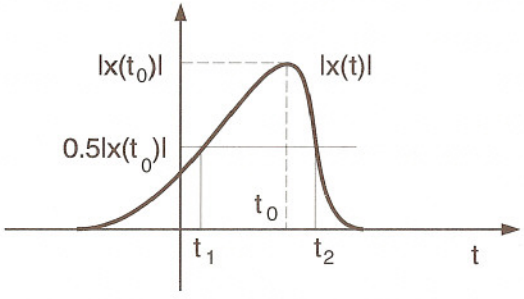


Figura 4.36 Definizione di durata di un segnale

Oltre alla banda a –3 dB (di gran lunga la più utilizzata) e alla durata a metà ampiezza, esistono altre definizioni di banda e durata particolarmente importanti dal punto di vista concettuale, come la *durata efficace* (o *quadratica*) e la *banda efficace* (*quadratica*). Cominciamo a esaminare la definizione di durata efficace per un segnale $x(t)$ reale e ad energia finita. Come nella definizione di durata a metà ampiezza, viene preliminarmente identificato un valore temporale di riferimento t_0 che in un certo senso rappresenta il "centro" dell'impulso:

$$t_0 \triangleq \frac{1}{E_x}\int_{-\infty}^{\infty} t\,x^2(t)\,dt \quad , \quad E_x=\int_{-\infty}^{\infty} x^2(t)\,dt \tag{4.3.21}$$

dove E_x è l'energia del segnale. La *durata efficace* D_q è allora definita dalla relazione

$$D_q^2 \triangleq \frac{1}{E_x} \int_{-\infty}^{\infty} (t - t_0)^2 x^2(t)\, dt \tag{4.3.22}$$

che fornisce un'indicazione della dispersione dei valori del segnale attorno al centro presunto t_0. Il concetto di durata efficace è analogo a quello di *varianza* di una densità di probabilità $|x(t)|^2 / E_x$ attorno al *valor medio* t_0. In maniera analoga, si definisce una *banda efficace* B_q come segue:

$$(2B_q)^2 \triangleq \frac{1}{E_x} \int_{-\infty}^{\infty} |X(f)|^2 f^2 df \tag{4.3.23}$$

in cui il "centro" (valor medio) dello spettro $|X(f)|^2$ non compare perché risulta comunque pari a 0. Il fattore 2 a moltiplicare la B_q (che non compare nella definizione di durata) deriva dal fatto che la banda viene come sempre misurata solo sul semiasse positivo delle frequenze, mentre il secondo membro dà una indicazione della dispersione dello spettro sia a *destra* che a *sinistra* della $f = 0$.

Una generalizzazione importante della proprietà già esaminata secondo la quale un segnale non può contemporaneamente presentare banda limitata e durata limitata è costituita dalla relazione seguente:

$$B_q \cdot D_q \geq \frac{1}{8\pi} \tag{4.3.24}$$

Questo risultato mostra come non si possa ridurre arbitrariamente la durata di un segnale senza far crescere la sua banda e viceversa: il prodotto durata-banda è comunque *limitato inferiormente*. Il particolare valore del limite inferiore, $1/8\pi$, deriva dalle particolari definizioni di banda e durata adottate, ma il principio è di carattere generale: con diverse definizioni di banda e durata si ottiene comunque un limite inferiore, ovviamente diverso da quello del caso presente, ma sempre finito.

Una relazione della forma (4.3.24) appare anche in meccanica quantistica con il nome di *principio di indeterminazione di Heisenberg*. Questa legge fisica fondamentale sancisce l'impossibilità di ridurre arbitrariamente *l'incertezza* sulla misurazione di variabili quantistiche coniugate (cioè la dispersione dei valori misurati), come la posizione e la quantità di moto di una particella.

Nel Paragrafo 4.4, dimostreremo che per un qualunque segnale reale a energia finita $y(t)$ si ha:

$$E_y = \int_{-\infty}^{\infty} y^2(t)\,dt = \int_{-\infty}^{\infty} |Y(f)|^2\,df \tag{4.3.25}$$

Questo risultato (teorema di Parseval (4.4.4)), insieme al teorema di derivazione del Paragrafo 3.3.6, è utile per riscrivere l' espressione della banda efficace (4.3.23) come segue:

$$B_q^2 = \frac{1}{16\pi^2 E_x} \int_{-\infty}^{\infty} |j2\pi f\, X(f)|^2\,df = \frac{1}{16\pi^2 E_x} \int_{-\infty}^{\infty} \left[\frac{dx(t)}{dt}\right]^2 dt \tag{4.3.26}$$

Supponiamo, senza perdita di generalità, che il centro del segnale t_0 sia pari a zero. Il prodotto fra D_q^2 e B_q^2 è allora espresso da (vedi la (4.3.22) e la (4.3.26))

$$D_q^2\, B_q^2 = \frac{1}{16\pi^2 E_x^2} \int_{-\infty}^{\infty} t^2 x^2(t)\,dt \int_{-\infty}^{\infty} \left[\frac{dx(t)}{dt}\right]^2 dt \tag{4.3.27}$$

Richiamiamo a questo punto la *disuguaglianza di Schwarz*:

$$\int_{-\infty}^{\infty} x_a^2(t)\,dt \int_{-\infty}^{\infty} x_b^2(t)\,dt \geq \left[\int_{-\infty}^{\infty} x_a(t)\, x_b(t)\,dt\right]^2 \tag{4.3.28}$$

valida per ogni coppia $x_a(t)$, $x_b(t)$ di funzioni reali a energia finita. Applicando la disuguaglianza al prodotto durata-banda, si può scrivere che

$$D_q^2\, B_q^2 \geq \frac{1}{16\pi^2 E_x^2} \left[\int_{-\infty}^{\infty} t\, x(t) \frac{dx(t)}{dt} dt\right]^2 \tag{4.3.29}$$

Calcolando per parti l'integrale a secondo membro si ha

$$\int_{-\infty}^{\infty} t\, x(t) \frac{dx(t)}{dt}\,dt = \frac{t\, x^2(t)}{2}\Bigg|_{-\infty}^{+\infty} - \frac{1}{2}\int_{-\infty}^{\infty} x^2(t)\,dt = -\frac{1}{2}E_x \tag{4.3.30}$$

avendo osservato che necessariamente, per un segnale a energia finita, $\lim_{t\to\pm\infty} t\, x^2(t) = 0$. Sostituendo la (4.3.30) nella (4.3.29) si ottiene infine

$$D_q^2\, B_q^2 \geq \frac{1}{64\pi^2} \quad \Rightarrow \quad D_q\, B_q \geq \frac{1}{8\pi} \tag{4.3.31}$$

La relazione è valida con il segno di uguaglianza quando la disuguaglianza di Schwarz (4.3.28) è anch'essa verificata con il segno di uguaglianza. Ciò accade quando il segnale $x_a(t)$ è *proporzionale* a $x_b(t)$, e quindi, nel nostro caso, quando (vedi la (4.3.29)) la derivata prima di $x(t)$ è proporzionale a $t\,x(t)$.

Esempio 4.10

Calcoliamo la banda e la durata efficace del segnale *esponenziale bilatero*:

$$x(t) = e^{-|t|/T} \tag{E4.10.1}$$

Poiché il segnale è pari, il suo centro t_0 espresso dalla (4.3.21) è nullo. Pertanto la sua *durata efficace* D_q è

$$D_q = \sqrt{\frac{1}{E_x}\int_{-\infty}^{\infty} t^2 x^2(t)\,dt} \tag{E4.10.2}$$

L'energia E_x del segnale $x(t)$ è data da

$$E_x = \int_{-\infty}^{\infty} x^2(t)\,dt = 2\int_0^{\infty} e^{-2t/T} dt = T \tag{E4.10.3}$$

e inoltre

$$\int_{-\infty}^{\infty} t^2 x^2(t)\,dt = 2\int_0^{\infty} t^2 e^{-2t/T}\,dt = \frac{T^3}{4}\int_0^{\infty} \alpha^2 e^{-\alpha}\,d\alpha = \frac{T^3}{2} \tag{E4.10.4}$$

Sostituendo, si ricava che

$$D_q = \frac{T}{\sqrt{2}} \tag{E4.10.5}$$

Calcoliamo ora la banda efficace attraverso la definizione modificata (4.3.26):

$$B_q = \frac{1}{4\pi}\sqrt{\frac{1}{E_x}\int_{-\infty}^{\infty}\left[\frac{dx(t)}{dt}\right]^2 dt} \tag{E4.10.6}$$

Poiché

$$\int_{-\infty}^{\infty}\left[\frac{dx(t)}{dt}\right]^2 dt = \frac{1}{T^2}\int_{-\infty}^{\infty} e^{-2t/T} dt = \frac{E_x}{T^2} \tag{E4.10.7}$$

dalla (E4.10.6) si ricava che

$$B_q = \frac{1}{4\pi T} \tag{E4.10.8}$$

Il prodotto fra durata e banda efficace del segnale (E4.10.1) è quindi dato da

$$B_q\, D_q = \frac{1}{4\pi\sqrt{2}} > \frac{1}{8\pi} \tag{E4.10.9}$$

❑

Esempio 4.11

Calcoliamo la banda e la durata efficace del segnale *Gaussiano*

$$x(t) = e^{-(t/T)^2} \tag{E4.11.1}$$

Poiché il segnale è pari, il suo centro t_0 è nullo e la sua *durata efficace* D_q è

$$D_q = \sqrt{\frac{1}{E_x}\int_{-\infty}^{\infty} t^2 x^2(t)\, dt} \tag{E4.11.2}$$

Dallo studio delle variabili aleatorie Gaussiane è noto che, assegnato un valore $\sigma > 0$, si ha

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-t^2/2\sigma^2}\, dt = 1 \tag{E4.11.3}$$

e

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} t^2 e^{-t^2/2\sigma^2}\, dt = \sigma^2 \tag{E4.11.4}$$

Dunque l'energia E_x del segnale $x(t)$ è pari a

$$E_x = \int_{-\infty}^{\infty} x^2(t)\, dt = \int_{-\infty}^{\infty} e^{-2(t/T)^2} dt = \sigma\sqrt{2\pi}\,\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-t^2/2\sigma^2}\, dt = \sigma\sqrt{2\pi} \tag{E4.11.5}$$

ove $\sigma = T/\sqrt{2}$. L'integrale che figura a secondo membro della (E4.11.2) vale

$$\int_{-\infty}^{\infty} t^2 x^2(t)\, dt = \int_{-\infty}^{\infty} t^2\, e^{-2(t/T)^2}\, dt = \sigma\sqrt{2\pi}\,\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} t^2 e^{-t^2/2\sigma^2}\, dt = \sigma^3\sqrt{2\pi} \tag{E4.11.6}$$

dalla (E4.10.6) si ricava che

$$B_q = \frac{1}{4\pi T} \tag{E4.10.8}$$

Il prodotto fra durata e banda efficace del segnale (E4.10.1) è quindi dato da

$$B_q\, D_q = \frac{1}{4\pi\sqrt{2}} > \frac{1}{8\pi} \tag{E4.10.9}$$

❑

Esempio 4.11

Calcoliamo la banda e la durata efficace del segnale *Gaussiano*

$$x(t) = e^{-(t/T)^2} \tag{E4.11.1}$$

Poiché il segnale è pari, il suo centro t_0 è nullo e la sua *durata efficace* D_q è

$$D_q = \sqrt{\frac{1}{E_x}\int_{-\infty}^{\infty} t^2 x^2(t)\, dt} \tag{E4.11.2}$$

Dallo studio delle variabili aleatorie Gaussiane è noto che, assegnato un valore $\sigma > 0$, si ha

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-t^2/2\sigma^2}\, dt = 1 \tag{E4.11.3}$$

e

$$\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} t^2 e^{-t^2/2\sigma^2}\, dt = \sigma^2 \tag{E4.11.4}$$

Dunque l'energia E_x del segnale $x(t)$ è pari a

$$E_x = \int_{-\infty}^{\infty} x^2(t)\, dt = \int_{-\infty}^{\infty} e^{-2(t/T)^2}\, dt = \sigma\sqrt{2\pi}\,\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{-t^2/2\sigma^2}\, dt = \sigma\sqrt{2\pi} \tag{E4.11.5}$$

ove $\sigma = T/\sqrt{2}$. L'integrale che figura a secondo membro della (E4.11.2) vale

$$\int_{-\infty}^{\infty} t^2 x^2(t)\, dt = \int_{-\infty}^{\infty} t^2\, e^{-2(t/T)^2}\, dt = \sigma\sqrt{2\pi}\,\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} t^2 e^{-t^2/2\sigma^2}\, dt = \sigma^3\sqrt{2\pi} \tag{E4.11.6}$$

e quindi

$$D_q^2 = \frac{\sigma^3\sqrt{2\pi}}{\sigma\sqrt{2\pi}} = \sigma^2 \quad \Rightarrow \quad D_q = \frac{T}{\sqrt{2}} \tag{E4.11.7}$$

Per calcolare la banda efficace utilizziamo di nuovo la (4.3.26)

$$B_q = \frac{1}{4\pi}\sqrt{\frac{1}{E_x}\int_{-\infty}^{\infty}\left[\frac{dx(t)}{dt}\right]^2 dt} \tag{E4.11.8}$$

Poiché

$$\frac{dx(t)}{dt} = -\frac{2t}{T^2}e^{-(t/T)^2} = -\frac{2}{T^2}t\,x(t) \tag{E4.11.9}$$

si ha

$$\int_{-\infty}^{\infty}\left[\frac{dx(t)}{dt}\right]^2 dt = \frac{4}{T^4}\int_{-\infty}^{\infty}t^2x^2(t)\,dt = \frac{1}{4\sigma^4}\sigma^3\sqrt{2\pi} = \frac{\sqrt{2\pi}}{4\sigma} \tag{E4.11.10}$$

per cui si ricava

$$B_q^2 = \frac{1}{16\pi^2\sigma\sqrt{2\pi}}\frac{\sqrt{2\pi}}{4\sigma} = \frac{1}{64\pi^2\sigma^2} \quad \Rightarrow \quad B_q = \frac{\sqrt{2}}{8\pi T} \tag{E4.11.11}$$

Il prodotto fra durata e banda efficace del segnale (E4.11.1) è quindi dato da

$$B_q\,D_q = \frac{1}{8\pi} \tag{E4.11.12}$$

Si osserva che in questo caso il prodotto banda-durata assume proprio *il minimo valore possibile*. Ciò accade perché la derivata prima del segnale $x(t)$ è proporzionale al segnale $t\,x(t)$, come mostra la (E4.11.9). ❑

4.3.4 Distorsioni introdotte dai filtri

Abbiamo già esaminato una funzione tipica di un *filtro*, cioè quella di separare un segnale utile da altri segnali di disturbo. In questa operazione, si deve porre la massima attenzione a non alterare o, più precisamente, a non *distorcere* il segnale utile per preservarne intatto il contenuto informativo. Per trovare dei criteri che rispondono a questa esigenza, cominciamo con lo stabilire sotto quali condizioni il segnale in uscita $y(t)$ da un generico SLS rappresenta una *replica*

fedele del segnale di ingresso $x(t)$. La definizione di replica fedele è in generale dipendente dall'applicazione; in molti casi però è utile dire che il SLS *non introduce distorsioni* quando l'uscita del sistema è

$$y(t) = K\, x(t - t_0) \tag{4.3.32}$$

ove K è una costante reale e t_0 è una traslazione temporale (ritardo) (Figura 4.37). Passando alle trasformate di Fourier otteniamo

$$Y(f) = K\, X(f)\, e^{-j2\pi f t_0} \tag{4.3.33}$$

da cui si trova la seguente espressione per la risposta in frequenza del sistema:

$$H(f) = \frac{Y(f)}{X(f)} = K\, e^{-j2\pi f t_0} \tag{4.3.34}$$

e per le corrispondenti risposte in ampiezza $A(f)$ e fase $\theta(f)$:

$$A(f) = |K| \quad , \quad \theta(f) = -2\pi f t_0 \tag{4.3.35}$$

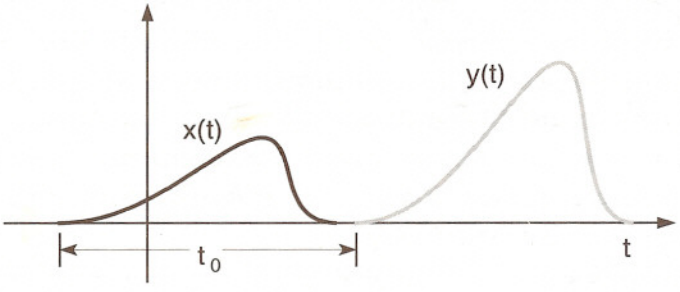


Figura 4.37 Il segnale $y(t)$ è una replica fedele di $x(t)$

Quindi, affinché un sistema non introduca distorsioni, esso deve possedere una *risposta in ampiezza costante* e una *risposta in fase proporzionale alla frequenza* come illustrato in Figura 4.38. In altri termini, le componenti sinusoidali in cui un segnale arbitrario può pensarsi scomposto devono essere amplificate o attenuate tutte nella medesima misura, e devono essere ritardate ciascuna della medesima quantità (e quindi *sfasate* di un angolo proporzionale alla frequenza della componente stessa). Tuttavia è chiaro che ogni sistema reale ha dei limiti di banda intrinseci e non può garantire una risposta in frequenza con le suddette caratteristiche per tutti i valori della frequenza. Sulla base di questa considera-

zione, sembrerebbe irrealizzabile la condizione di non distorsione appena discussa. Inoltre, uno qualunque dei filtri già esaminati nelle pagine precedenti parrebbe, in base a questa discussione, un sistema *distorcente* perché per definizione esso non possiede, in particolare, risposta in ampiezza costante.

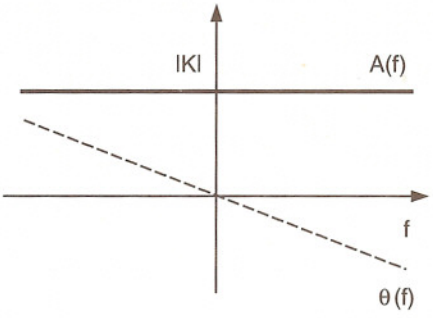


Figura 4.38 Risposta in ampiezza e in fase di un SLS che non introduce distorsioni

In realtà, il segnale "utile" $x(t)$ che non deve essere distorto sarà in generale caratterizzato da una banda limitata; le condizioni (4.3.35) possono allora essere verificate *soltanto per tutte le frequenze all'interno della banda del segnale.* Ai fini della distorsione del segnale non ha cioè rilevanza l'andamento delle risposte del sistema per frequenze in corrispondenza alle quali non esistono componenti nello spettro del segnale stesso. Se, ad esempio, la trasformata $X(f)$ e le risposte in ampiezza e fase del sistema sono quelle rappresentate in Figura 4.39, allora il segnale in uscita *è una replica indistorta di* $x(t)$ anche se il filtro non ha risposte come in Figura 4.38.

Se non si riesce a garantire le condizioni di non-distorsione neanche nella banda del segnale, questo subisce *distorsioni lineari.* In particolare, se la risposta in ampiezza non è costante nella banda del segnale, si avranno *distorsioni di ampiezza* (Figura 4.40a); viceversa, se la risposta in fase non è lineare nella banda del segnale, si avranno *distorsioni di fase* (Figura 4.40b). Per alcuni tipi di segnale, è tollerabile l'introduzione di distorsioni di fase da parte di un filtro. Ad esempio, nei segnali telefonici di tipo vocale, la presenza di una distorsione di fase introdotta dal sistema di comunicazione è praticamente ininfluente sulla qualità del segnale ricevuto. Viceversa la distorsione di fase non è tollerabile per i segnali di tipo numerico delle trasmissioni facsimile che peraltro transitano attraverso gli stessi canali telefonici. I ricevitori negli apparecchi fax possiedono degli appositi circuiti cosiddetti *equalizzatori* il cui compito è proprio quello di annullare le eventuali distorsioni di fase introdotte dai circuiti telefonici ordinari.

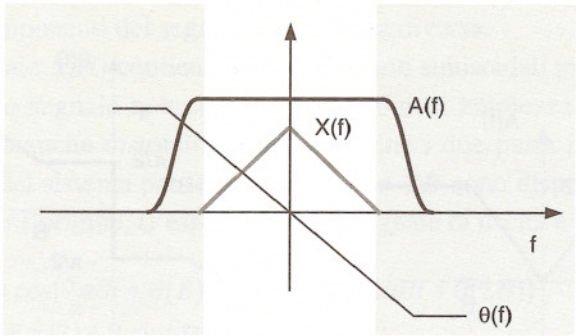


Figura 4.39 Il sistema non introduce distorsioni sul segnale $x(t)$

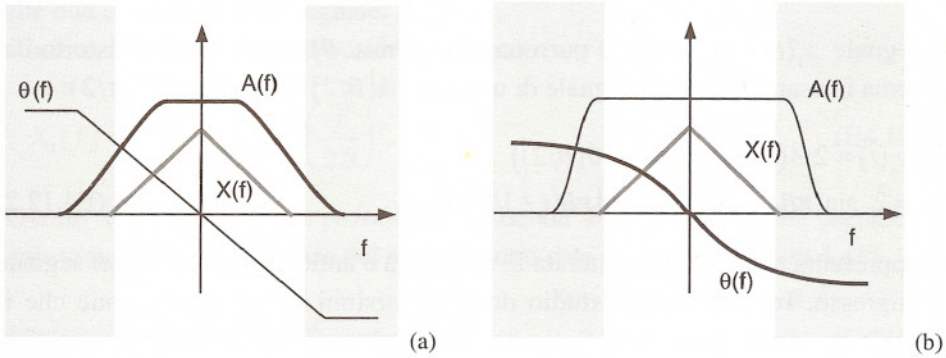


Figura 4.40 Distorsioni lineari di ampiezza (a) e fase (b)

Esempio 4.12

Consideriamo i seguenti quattro segnali:

$$x_1(t) = 2\sin(\pi Bt) \tag{E4.12.1a}$$

$$x_2(t) = \cos(\pi Bt) + \sin(3\pi Bt) \tag{E4.12.1b}$$

$$x_3(t) = 4\cos(2\pi Bt) + \sin(4\pi Bt) \tag{E4.12.1c}$$

$$x_4(t) = \mathrm{sinc}^2(Bt) \tag{E4.12.1d}$$

e cerchiamo di stabilire se tali segnali vengono o meno distorti nel passaggio

attraverso il SLS la cui risposta in frequenza è rappresentata in Figura 4.41.

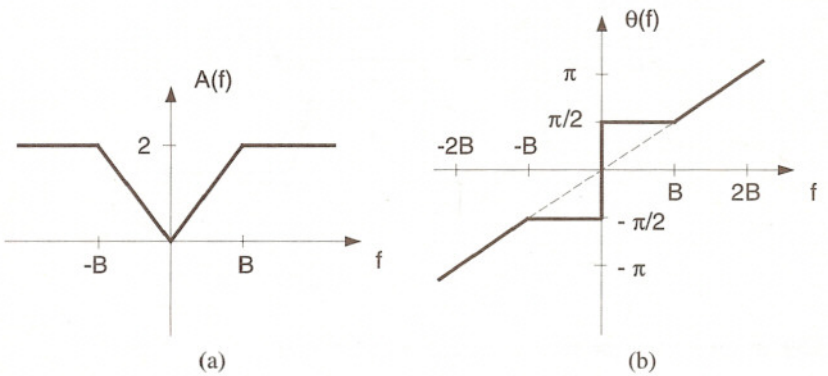


Figura 4.41 Risposta in ampiezza (a) e in fase (b) del SLS dell'Esempio 4.12

Il segnale $x_1(t)$ è sinusoidale puro con frequenza $B/2$ e non viene distorto dal sistema in esame. Infatti il segnale di uscita è ($A(B/2)=1$ e $\theta(B/2)=\pi/2$):

$$\begin{aligned} y_1(t) &= 2\,A(B/2)\sin\big(\pi Bt+\theta(B/2)\big) \\ &= 2\,\sin(\pi Bt+\pi/2)=2\sin\big[\pi B(t+1/2B)\big] \end{aligned} \qquad \text{(E4.12.2)}$$

e rappresenta una replica, inalterata in ampiezza e anticipata di $1/2B$, del segnale di ingresso. In generale, lo studio delle distorsioni lineari presuppone che il segnale in ingresso a un SLS sia costituito almeno da due componenti frequenziali che possono subire, nell'attraversare il sistema, ritardi e amplificazioni diverse. Invece, in presenza di una sola componente frequenziale, come in questo caso, il segnale di uscita sarà sempre una replica indistorta del segnale di ingresso comunque esso venga amplificato e/o ritardato.

Il segnale $x_2(t)$ contiene invece due componenti sinusoidali alle due diverse frequenze $B/2$ e $3B/2$. Questo segnale subisce distorsioni di ampiezza poiché $A(B/2)=1\neq A(3B/2)=2$. Inoltre esso subisce distorsioni di fase poiché i due punti individuati sulla risposta in fase del sistema per le frequenze $B/2$ e $3B/2$ non sono disposti secondo una retta passante per l'origine degli assi cartesiani. La presenza di distorsioni può anche essere messa in evidenza scrivendo l'espressione del segnale di uscita

$$\begin{aligned} y_2(t) &= A(B/2)\cos\big(\pi Bt+\theta(B/2)\big)+A(3B/2)\sin\big(3\pi Bt+\theta(3B/2)\big) \\ &= \cos(\pi Bt+\pi/2)+2\sin(3\pi Bt+3\pi/4) \\ &= \cos\big[\pi B(t+1/2B)\big]+2\sin\big[3\pi B(t+1/4B)\big] \end{aligned} \qquad \text{(E4.12.3)}$$

La presenza di distorsioni di fase si traduce in un anticipo diverso introdotto dal sistema sulle componenti del segnale a frequenza diversa.

Anche il segnale $x_3(t)$ contiene due componenti sinusoidali ma alle frequenze B e $2B$. Questo segnale non subisce distorsioni di ampiezza poiché $A(B) = A(2B) = 1$, e neanche distorsioni di fase poiché i due punti individuati sulla risposta in fase del sistema per le frequenze B e $2B$ sono disposti secondo una retta passante per l'origine. L'espressione del segnale di uscita è

$$\begin{aligned} y_3(t) &= 4A(B)\cos\big(2\pi Bt + \theta(B)\big) + A(2B)\sin\big(4\pi Bt + \theta(2B)\big) \\ &= 8\cos\big(2\pi Bt + \pi/2\big) + 2\sin(4\pi Bt + \pi) \\ &= 8\cos\big[2\pi B(t + 1/4B)\big] + 2\sin\big[4\pi B(t + 1/4B)\big] \end{aligned} \qquad \text{(E4.12.4)}$$

e mostra che il sistema introduce la stessa amplificazione e lo stesso anticipo sulle due componenti del segnale.

Infine, la trasformata di Fourier del segnale $x_4(t)$ è

$$X_4(f) = \frac{1}{B}\left(1 - \frac{|f|}{B}\right)\mathrm{rect}\left(\frac{f}{2B}\right) \qquad \text{(E4.12.5)}$$

Pertanto $x_4(t)$ è un segnale passa-basso di banda B. Nella banda del segnale, le risposte in ampiezza e in fase del sistema sono date da (vedi la Figura 4.41)

$$A(f) = \frac{2|f|}{B} \quad , \quad \theta(f) = \frac{\pi}{2}\,\mathrm{sgn}(f) \qquad \text{(E4.12.6)}$$

e cioè

$$H(f) = A(f)\,e^{j\theta(f)} = j\frac{2f}{B} = \frac{1}{\pi B}(j2\pi f) \quad , \quad -B \le f \le B \qquad \text{(E4.12.7)}$$

Il segnale subisce ovviamente distorsioni di ampiezza e di fase. Si può anche osservare che il sistema, *nell'intervallo di frequenze di interesse*, esegue la *derivata* del segnale di ingresso (a meno di una costante $1/\pi B$). Allora l'espressione del segnale di uscita è

$$\begin{aligned} y_4(t) &= \frac{1}{\pi B}\frac{d}{dt}x_4(t) = \frac{1}{\pi B}\frac{d}{dt}\mathrm{sinc}^2(Bt) \\ &= \frac{2}{\pi B}\,\mathrm{sinc}(Bt)\frac{d}{dt}\mathrm{sinc}(Bt) = \frac{2}{\pi Bt}\,\mathrm{sinc}(Bt)\big[\cos(\pi Bt) - \mathrm{sinc}(Bt)\big] \end{aligned} \qquad \text{(E4.12.8)}$$

❑

I criteri di fedeltà, distorsione e banda appena introdotti possono essere rivisitati e ridefiniti considerando il contenuto *energetico* dei segnali. È quindi necessario esaminare più da vicino il concetto di *energia* e *potenza* dei segnali a tempo continuo, e di studiare come queste grandezze vengono influenzate dall'elaborazione dei medesimi con sistemi lineari stazionari.

4.4 Densità spettrale di energia e potenza

4.4.1 Teorema di Parseval e densità spettrale di energia

Consideriamo un segnale $x(t)$ a energia finita:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 \, dt < \infty \tag{4.4.1}$$

e mostriamo come il calcolo di questa grandezza può essere effettuato anche nel dominio della frequenza. Infatti

$$\int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} x(t)\, x^*(t)\, dt = \int_{t=-\infty}^{\infty} x(t) \left[\int_{f=-\infty}^{\infty} X^*(f)\, e^{-j2\pi ft} df \right] dt \tag{4.4.2}$$

Se a secondo membro della (4.4.2) si inverte l'ordine di integrazione si ottiene

$$\int_{f=-\infty}^{\infty} X^*(f) \left[\int_{t=-\infty}^{\infty} x(t)\, e^{-j2\pi ft} dt \right] df = \int_{f=-\infty}^{\infty} X^*(f)\, X(f)\, df = \int_{-\infty}^{\infty} |X(f)|^2 \, df \tag{4.4.3}$$

per cui si conclude:

$$\int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} |X(f)|^2 \, df \tag{4.4.4}$$

risultato detto *teorema (o relazione) di Parseval*. Al di là dell'utilità come strumento di calcolo, il teorema di Parseval costituisce la base per considerazioni di carattere più profondo. Dalla definizione di energia di un segnale data nel Capitolo 1, sappiamo che la quantità $p_x(t) = |x(t)|^2$ rappresenta la *potenza istantanea* (normalizzata) del segnale che, integrata, fornisce appunto l'energia totale del segnale stesso. Non è chiaro invece il significato fisico della quantità "duale"

$$\mathcal{E}_x(f) \triangleq |X(f)|^2 \tag{4.4.5}$$

Considerando $\mathcal{E}_x(f)$ come una funzione della frequenza, è chiaro che

- $\mathcal{E}_x(f)$ è una funzione che assume sempre valori non negativi:

$$\mathcal{E}_x(f) \geq 0 \tag{4.4.6}$$

- se il segnale $x(t)$ è reale, allora $\mathcal{E}_x(f)$ è una funzione pari:

$$\mathcal{E}_x(f) = \mathcal{E}_x(-f) \tag{4.4.7}$$

- l'energia del segnale $x(t)$ può essere calcolata integrando la funzione $\mathcal{E}_x(f)$ su tutto l'asse delle frequenze:

$$E_x = \int_{-\infty}^{\infty} \mathcal{E}_x(f)\, df \tag{4.4.8}$$

Queste *proprietà* non rispondono però alla domanda circa il *significato* da attribuire a $\mathcal{E}_x(f)$. Consideriamo allora un sistema lineare stazionario con risposta in frequenza $H(f)$ il cui ingresso sia $x(t)$ e la cui uscita sia $y(t)$. La funzione $\mathcal{E}_y(f)$ relativa al segnale di *uscita* si può ricavare immediatamente nota $\mathcal{E}_x(f)$:

$$\mathcal{E}_y(f) = |Y(f)|^2 = |X(f)\,H(f)|^2 = |X(f)|^2 |H(f)|^2 = \mathcal{E}_x(f)\,|H(f)|^2 \tag{4.4.9}$$

Stante questa relazione di carattere generale, scegliamo un *particolare* SLS per compiere una sorta di "esperimento ideale", e cioè il *filtro passa-banda ideale* di Figura 4.42.

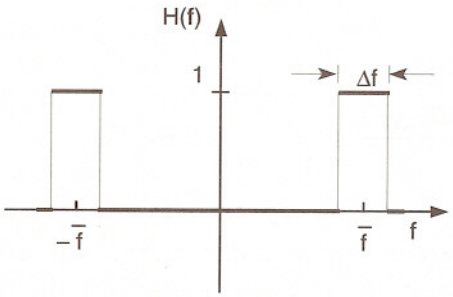


Figura 4.42 Filtro passa-banda ideale con frequenza centrale $\bar{f}$ e banda Δf

La funzione $\mathcal{E}_x(f)$ del segnale di ingresso è quella rappresentata a tratto sottile in Figura 4.43, mentre l'andamento della corrispondente $\mathcal{E}_y(f)$ è disegnato a tratto spesso. L'energia del segnale di uscita $y(t)$ si può calcolare come segue:

$$E_y = \int_{-\infty}^{\infty} \mathcal{E}_y(f)\, df = \int_{-\infty}^{\infty} \mathcal{E}_x(f)\, |H(f)|^2\, df = \int_{-\bar{f}-\Delta f/2}^{-\bar{f}+\Delta f/2} \mathcal{E}_x(f)\, df + \int_{\bar{f}-\Delta f/2}^{\bar{f}+\Delta f/2} \mathcal{E}_x(f)\, df$$

$$= 2 \int_{\bar{f}-\Delta f/2}^{\bar{f}+\Delta f/2} \mathcal{E}_x(f)\, df \qquad (4.4.10)$$

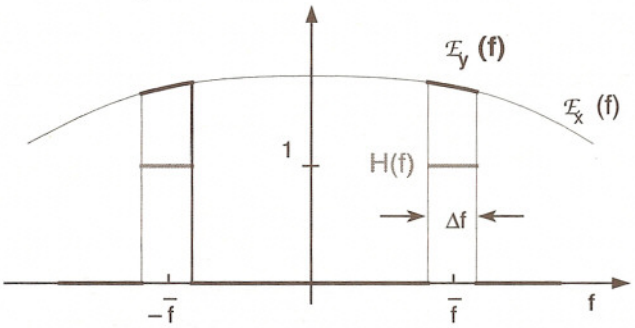


Figura 4.43 Andamento delle funzioni $\mathcal{E}_x(f)$ ed $\mathcal{E}_y(f)$

Se si riduce progressivamente la banda passante Δf del filtro, cioè si considera un filtro estremamente *selettivo*, si può approssimare la funzione $\mathcal{E}_x(f)$ all'interno della banda stessa con una *costante* di valore pari a $\mathcal{E}_x(\bar{f})$. Allora, sotto questa ipotesi, l'energia del segnale d'uscita può essere approssimata da

$$E_y \cong 2\Delta f \cdot \mathcal{E}_x(\bar{f}) \qquad (4.4.11)$$

e quindi

$$\mathcal{E}_x(\bar{f}) \cong \frac{1}{2}\frac{E_y}{\Delta f} = \frac{1}{2}\frac{\Delta E_x(\bar{f})}{\Delta f} \qquad (4.4.12)$$

Notiamo, ora, che $y(t)$ è stato ottenuto da $x(t)$ sopprimendo tutte le componenti frequenziali al di fuori di un intorno della frequenza $\pm\bar{f}$. Dunque, come suggerisce la notazione nella (4.4.12), $E_y = \Delta E_x(\bar{f})$ rappresenta il *contributo* all'energia totale del segnale E_x delle sole componenti di $x(t)$ con frequenze appartenenti all'intervallo $[\bar{f} - \Delta f/2, \bar{f} + \Delta f/2\,]$ (più il simmetrico sul semiasse negativo), in quanto tutte le altre componenti sono state *cancellate* dal filtro passa-banda.

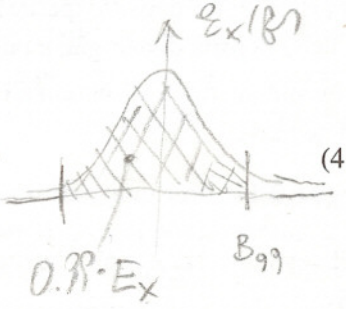
Questo risultato permette finalmente di fornire un'interpretazione della funzione $\mathcal{E}_x$: il particolare valore $\mathcal{E}_x(\bar{f})$ alla frequenza $\bar{f}$ rappresenta il contributo *locale* all'energia totale del segnale $x(t)$ dovuto alle *sole* componenti di segnale

con frequenza appartenente a un piccolo intorno di $\bar{f}$, rapportato all'ampiezza piccola Δf dell'intorno stesso. Ciò corrisponde alla classica definizione di *densità* di una grandezza fisica (qui l'energia) rispetto a una misura di estensione (qui l'ampiezza di una banda), e giustifica il nome con il quale si designa la funzione $\mathcal{E}_x(f)$: *densità spettrale di energia* del segnale $x(t)$. Il fattore 1/2 nella (4.4.12) è relativo alla necessità di considerare sia le componenti in prossimità di $\bar{f}$ sia le "omologhe" a frequenza $-\bar{f}$ in modo da dar luogo a oscillazioni reali.

La definizione di densità spettrale di energia $\mathcal{E}_x(f) = |X(f)|^2$ mostra che, ai fini del calcolo dell'energia di $x(t)$, risulta determinante il solo *spettro di ampiezza* del segnale stesso, mentre lo spettro di fase è *ininfluente*. La stessa osservazione vale anche a proposito della relazione fondamentale del filtraggio $\mathcal{E}_y(f) = \mathcal{E}_x(f)|H(f)|^2$ secondo la quale, ai fini del calcolo dell'energia del segnale di uscita, risulta determinante la sola *risposta in ampiezza* del sistema, mentre la sua risposta in fase è del tutto ininfluente. In particolare, se vogliamo che l'energia del segnale d'ingresso *si conservi* nel filtraggio con un SLS, è necessario garantire che la sua risposta in ampiezza non cancelli quelle componenti alle frequenze per le quali $\mathcal{E}_x(f)$ assume i valori più significativi.

Per un segnale passa-basso si adotta talvolta una definizione di banda basata su un criterio *energetico*. La cosiddetta *banda al 99% dell'energia* del segnale $x(t)$ è il limite di banda B_{99} di quel filtro passa-basso ideale che, applicato sul segnale in oggetto, rende in uscita un segnale avente un'energia pari al 99% dell'energia d'ingresso E_x:

$$\int_{-B_{99}}^{B_{99}} \mathcal{E}_x(f)df = 0.99E_x \Rightarrow \frac{\int_{-B_{99}}^{B_{99}} \mathcal{E}_x(f)df}{\int_{-\infty}^{\infty} \mathcal{E}_x(f)df} = 0.99 \qquad (4.4.13)$$



4.4.2 Densità spettrale di potenza

I concetti appena introdotti riguardo ai segnali a energia finita possono essere generalizzati al caso di un segnale $x(t)$ per il quale è finita la *potenza*:

$$P_x = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t)\,dt < \infty \qquad (4.4.14)$$

Ricordiamo (si veda il Paragrafo 1.3) che il valore dell'integrale nella defini-

zione di potenza (4.4.14) rappresenta *l'energia* del segnale "troncato" $x_T(t) = x(t)\,\mathrm{rect}(t/T)$, che ha necessariamente energia finita. La densità spettrale di energia di $x_T(t)$ è dunque

$$\mathcal{E}_{x_T}(f) = |X_T(f)|^2 \qquad \text{d.d.e.} \tag{4.4.15}$$

e l'energia associata a $x_T(t)$ è data da

$$E_{x_T} = \int_{-\infty}^{\infty} |X_T(f)|^2 \, df \tag{4.4.16}$$

Allora la potenza del segnale di partenza $x(t)$ può essere calcolata come

$$P_x = \lim_{T\to\infty} \frac{E_{x_T}}{T} = \lim_{T\to\infty} \int_{-\infty}^{\infty} \frac{|X_T(f)|^2}{T}\, df = \int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{|X_T(f)|^2}{T}\, df \tag{4.4.17}$$

Questo risultato suggerisce di definire una funzione *densità spettrale di potenza* come limite della densità di energia del segnale troncato sull'intervallo "di osservazione" $[-T/2, T/2]$, rapportata all'ampiezza T dell'intervallo medesimo:

$$S_x(f) \triangleq \lim_{T\to\infty} \frac{\mathcal{E}_{x_T}(f)}{T} = \lim_{T\to\infty} \frac{|X_T(f)|^2}{T} \qquad \text{d.d.p.} \tag{4.4.18}$$

Le proprietà della densità spettrale *di potenza* sono formalmente analoghe a quelle della densità di energia, e possono essere facilmente dimostrate:

- la densità spettrale di potenza è una funzione della frequenza reale a valori non negativi:

$$S_x(f) \geq 0 \tag{4.4.19}$$

- se il segnale $x(t)$ è reale, la sua densità spettrale di potenza è una funzione pari:

$$S_x(f) = S_x(-f) \tag{4.4.20}$$

- la potenza di un segnale si può calcolare integrando la sua densità spettrale di potenza su tutto l'asse delle frequenze:

$$P_x = \int_{-\infty}^{\infty} S_x(f)\, df \tag{4.4.21}$$

- se il segnale $x(t)$ è posto in ingresso a un SLS con risposta in frequenza $H(f)$, la densità spettrale di potenza $S_y(f)$ del segnale di uscita $y(t)$ è legata a quella del segnale di ingresso dalla relazione:

$$S_y(f) = S_x(f)\,|H(f)|^2 \tag{4.4.22}$$

Anche il significato fisico della densità spettrale di potenza può essere compreso ricordando le considerazioni esposte a proposito della densità spettrale di energia; scelto un valore arbitrario $\bar{f}$ di frequenza, la quantità infinitesima $S_x(\bar{f})\,df$ rappresenta il contributo alla *potenza* totale del segnale $x(t)$ fornito dalle componenti frequenziali prossime a $\bar{f}$.

4.4.3 Funzione di autocorrelazione e teorema di Wiener-Khintchine

Le funzioni densità spettrale di energia e di potenza definite precedentemente possono essere calcolate in una maniera alternativa rispetto a quanto previsto dalle relative definizioni se si introduce una nuova grandezza che caratterizza ulteriormente il segnale nel dominio del tempo: la *funzione di autocorrelazione.* Per un segnale $x(t)$ a energia finita questa grandezza è definita da

$$R_x(\tau) \triangleq \int_{-\infty}^{\infty} x(t)\,x(t-\tau)\,dt \tag{4.4.23}$$

Il calcolo della funzione di autocorrelazione richiede di considerare il segnale dato $x(t)$, una sua replica $x(t-\tau)$ *ritardata del ritardo* τ, eseguire il prodotto di queste due funzioni e calcolare infine il valore dell'integrale del prodotto. Una volta eseguito il calcolo dell'integrale, il risultato resta ovviamente dipendente dal *ritardo* della replica $x(t-\tau)$ del segnale dato: la notazione $R_x(\tau)$ suggerisce proprio questa dipendenza.

La funzione di autocorrelazione fornisce informazioni utili sulla *rapidità di variazione* del segnale $x(t)$, come si può capire dall'esempio illustrato in Figura 4.44. Essa riassume infatti il calcolo del valore della funzione di autocorrelazione, per il medesimo valore di ritardo τ, di due differenti segnali, $x(t)$ e $y(t)$ aventi stessa energia ma due diverse "velocità di variazione". È chiaro che il segnale $x(t)$, più "veloce", presenta un valore della funzione di autocorrelazione minore di quello relativo al segnale $y(t)$ *a parità di ritardo* τ. Inoltre, quando la variabile τ assume valori crescenti si riduce l'ampiezza dell'intervallo in cui sia $x(t)$ che $x(t-\tau)$ assumono valori non nulli e, di conseguenza, tende a diminuire il valore di $R_x(\tau)$. Il nome di "autocorrelazione" suggerisce infatti che questa funzione è un indice di "somiglianza" del segnale con se stesso, o meglio, con

una replica di se stesso ritardata.

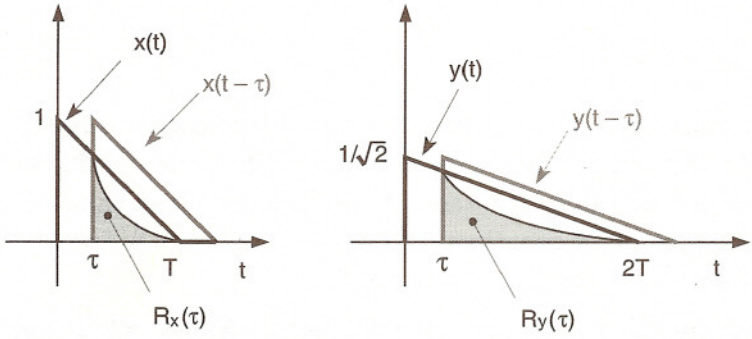


Figura 4.44 Esempi di calcolo della funzione di autocorrelazione

La funzione di autocorrelazione valutata per $\tau = 0$ coincide con l'energia del segnale $x(t)$:

$$R_x(0) = \int_{-\infty}^{\infty} x^2(t)\, dt = E_x \tag{4.4.24}$$

Inoltre, si intuisce facilmente che il segnale $x(t)$ è massimamente correlato (somigliante) con se stesso quando $\tau = 0$, quando cioè $x(t-\tau) = x(t)$. Possiamo quindi concludere (vedi l'Esercizio proposto 4.2) che

$$|R_x(\tau)| \leq R_x(0) \tag{4.4.25}$$

Il lettore dimostri anche che la $R_x(\tau)$ è una funzione *pari* del proprio argomento:

$$R_x(\tau) = R_x(-\tau) \tag{4.4.26}$$

Come conclusione di queste osservazioni, una forma verosimile per le funzioni di autocorrelazione dei due segnali di Figura 4.44 è quella di Figura 4.45. Quando la variabile τ supera una soglia di valore opportuno (detto *tempo di correlazione*) la funzione di autocorrelazione può considerarsi nulla.

Tornando all'argomento originario, è possibile mettere in relazione la funzione di autocorrelazione con la densità spettrale di energia del segnale; il *teorema di Wiener-Khintchine* afferma infatti che la *densità spettrale di energia di un segnale è pari alla trasformata di Fourier della rispettiva funzione di autocorrelazione*:

$$\mathcal{E}_x(f) = \int_{-\infty}^{\infty} R_x(\tau)\, e^{-j2\pi f\tau} d\tau = 2\int_{0}^{\infty} R_x(\tau)\cos(2\pi f\tau) d\tau \tag{4.4.27}$$

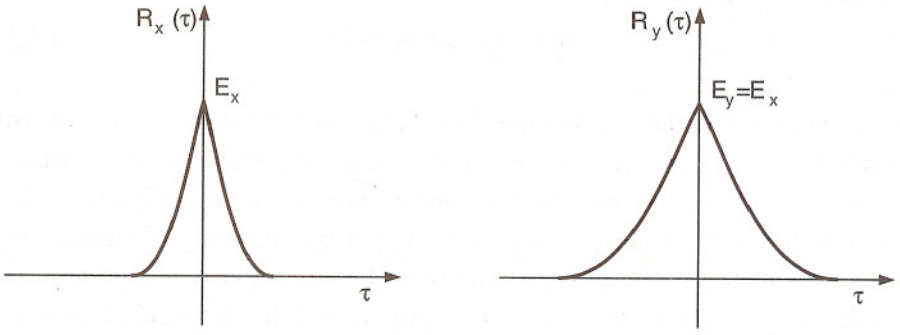


Figura 4.45 Funzioni di autocorrelazione dei segnali di Figura 4.44

Osserviamo infatti che la funzione di autocorrelazione può essere espressa nella forma di prodotto di convoluzione come segue:

$$R_x(\tau) = \int_{-\infty}^{\infty} x(\alpha)\, x(\alpha-\tau)\, d\alpha = \int_{-\infty}^{\infty} x(\alpha)\, x(-(\tau-\alpha))\, d\alpha = x(\tau)\otimes x(-\tau) \tag{4.4.28}$$

La trasformata di Fourier di $R_x(\tau)$ è allora

$$R_x(\tau) \Leftrightarrow X(f)\, X(-f) \tag{4.4.29}$$

ed essendo il segnale reale risulta: $X(-f) = X^*(f)$. Segue che

$$R_x(\tau) \Leftrightarrow |X(f)|^2 = \mathcal{E}_x(f) \tag{4.4.30}$$

La definizione di funzione di autocorrelazione data nella (4.4.23) si applica solo al caso di segnali a *energia finita*; per i segnali a potenza finita tale definizione deve essere modificata come segue:

$$R_x(\tau) \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x(t-\tau)\, dt \tag{4.4.31}$$

da cui, in particolare, $R_x(0) = P_x$. Questa funzione di autocorrelazione possiede tutte le proprietà formali già viste per i segnali a energia finita. Inoltre, il teorema di Wiener-Khintchine è di nuovo applicabile (si omette la dimostrazione,

formalmente più complessa di quella appena riportata per i segnali a energia finita), e stabilisce l'uguaglianza tra la trasformata di Fourier della funzione di autocorrelazione così definita e la densità spettrale *di potenza* del segnale $x(t)$:

$$S_x(f) = \int_{-\infty}^{\infty} R_x(\tau)\, e^{-j2\pi f\tau} d\tau = 2\int_{0}^{\infty} R_x(\tau)\cos(2\pi f\tau) d\tau \tag{4.4.32}$$

L'importanza del teorema di Wiener-Khintchine risiede nel fatto che la densità spettrale di energia (o potenza) di *ogni tipo di segnale* a energia o potenza finita, continuo o discreto, determinato o aleatorio, può essere ricavata attraverso il calcolo della trasformata di Fourier della opportuna funzione di autocorrelazione, come avremo modo di approfondire anche nel Capitolo 8. Questo procedimento si rivela spesso più comodo da utilizzare della definizione diretta di densità spettrale.

Esempio 4.13

Calcoliamo la densità spettrale di energia dell'impulso rettangolare

$$x(t) = \text{rect}\left(\frac{t}{T}\right) \tag{E4.13.1}$$

Essendo

$$X(f) = T\,\text{sinc}(fT) \tag{E4.13.2}$$

la densità spettrale di energia risulta

$$\mathcal{E}_x(f) = |X(f)|^2 = T^2\,\text{sinc}^2(fT) \tag{E4.13.3}$$

Allo stesso risultato si perviene calcolando la funzione di autocorrelazione che è data da

$$R_x(\tau) = x(\tau) \otimes x(-\tau) = \text{rect}\left(\frac{\tau}{T}\right) \otimes \text{rect}\left(\frac{\tau}{T}\right) = T\left(1 - \frac{|\tau|}{T}\right)\text{rect}\left(\frac{\tau}{2T}\right) \tag{E4.13.4}$$

Se si calcola la trasformata di Fourier della $R_x(\tau)$ si riottiene immediatamente il risultato (E4.13.3). ❑

Esempio 4.14

Calcoliamo la densità spettrale di potenza del segnale *gradino unitario*:

$$x(t) = \mathrm{u}(t) \tag{E4.14.1}$$

Questo segnale è aperiodico a potenza finita. La sua densità spettrale di potenza è espressa dalla (4.4.18):

$$S_x(f) = \lim_{T\to\infty} \frac{|X_T(f)|^2}{T} \tag{E4.14.2}$$

ove $X_T(f)$ è la trasformata di Fourier del segnale

$$x_T(t) = x(t)\,\mathrm{rect}\left(\frac{t}{T}\right) = \mathrm{rect}\left(\frac{t - T/4}{T/2}\right) \tag{E4.14.3}$$

come illustrato in Figura 4.46. Passando alle trasformate,

$$X_T(f) = \frac{T}{2}\,\mathrm{sinc}\left(\frac{fT}{2}\right) e^{-j\pi fT/2} \tag{E4.14.4}$$

e quindi la densità spettrale di potenza è data dal seguente limite:

$$S_x(f) = \lim_{T\to\infty} \frac{T}{4}\,\mathrm{sinc}^2\left(\frac{fT}{2}\right) = \lim_{T\to\infty} Y_T(f) \tag{E4.14.5}$$

con

$$Y_T(f) = \frac{T}{4}\,\mathrm{sinc}^2\left(\frac{fT}{2}\right) \tag{E4.14.6}$$

L'andamento della funzione $Y_T(f)$ è illustrato in Figura 4.47 per valori di T crescenti. Essa assume il valore $T/4$ per $f = 0$ e si annulla per le frequenze $f_k = 2k/T$, con k numero intero relativo diverso da 0. È importante notare che tale funzione *sottende un'area costante*, indipendentemente dal valore di T; infatti

$$\int_{-\infty}^{+\infty} Y_T(f)\,df = y_T(t)\big|_{t=0} = \frac{1}{2} \tag{E4.14.7}$$

in quanto

$$y_T(t) = \frac{1}{2}\left(1 - \frac{|t|}{T/2}\right) \text{rect}\left(\frac{t}{T}\right) \qquad \text{(E4.14.8)}$$

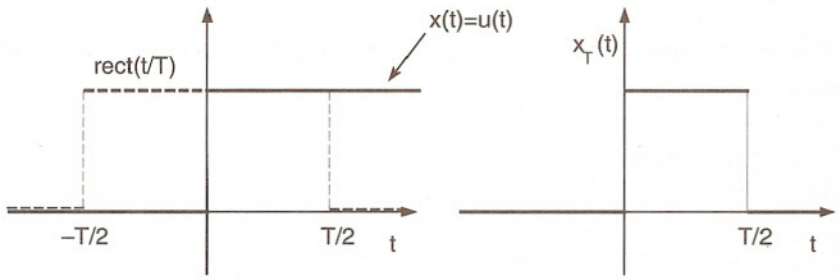


Figura 4.46 Andamento del segnale $x_T(t)$ (Esempio 4.14)

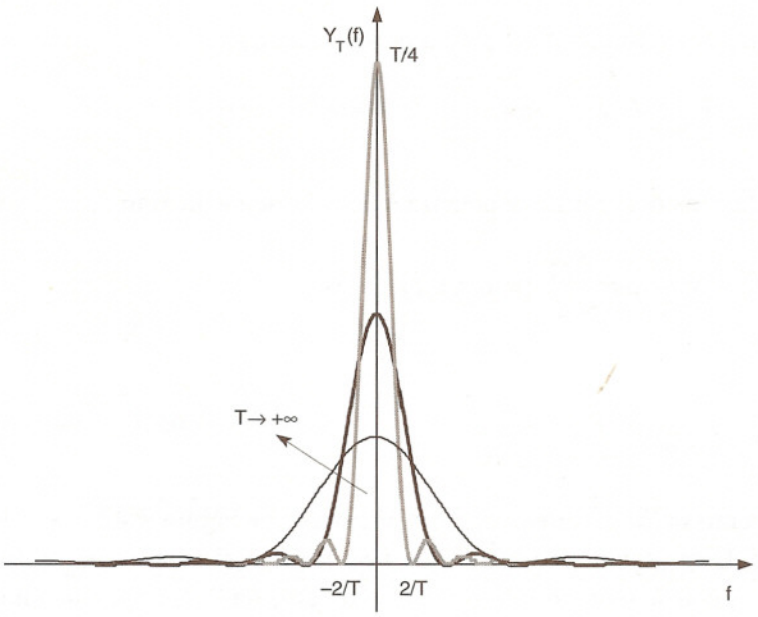


Figura 4.47 Andamento della funzione $Y_T(f)$ (Esempio 4.14)

Dunque, al crescere del valore di T il valore della funzione nell'origine *tende all'infinito*, la sua banda diminuisce arbitrariamente, poiché i suoi punti di nullo si avvicinano all'origine, ma l'area da essa sottesa non cambia: la successione di funzioni $Y_T(f)$ tende a un impulso di Dirac! In conclusione si ha:

$$S_x(f) = \frac{1}{2}\,\delta(f) \tag{E4.14.9}$$

Allo stesso risultato si perviene (tutto sommato in maniera più rapida) applicando il teorema di Wiener-Khintchine. La funzione di autocorrelazione per segnali a potenza finita è

$$R_x(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x(t-\tau)\, dt \tag{E4.14.10}$$

Limitiamoci al caso $\tau \geq 0$, visto che la funzione di autocorrelazione è pari. Allora si ottiene

$$R_x(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} \mathrm{u}(t)\, \mathrm{u}(t-\tau)\, dt = \lim_{T\to\infty} \frac{1}{T} \int_{0}^{T/2} \mathrm{u}(t-\tau)\, dt \tag{E4.14.11}$$

Calcoliamo adesso il valore dell'integrale, prima di eseguire il limite:

$$\int_0^{T/2} \mathrm{u}(t-\tau)\, dt = \left.\begin{cases} 0 & \tau > T/2 \\ T/2 - \tau & \tau \leq T/2 \end{cases}\right\} = (T/2 - \tau)\,\mathrm{u}(T/2 - \tau) \tag{E4.14.12}$$

quindi, per $\tau \geq 0$

$$\begin{aligned} R_x(\tau) &= \lim_{T\to\infty} \frac{1}{T}[(T/2 - \tau)u(T/2 - \tau)] \\ &= \lim_{T\to\infty} \left(\frac{1}{2} - \frac{\tau}{T}\right) u(T/2 - \tau) = \frac{1}{2} \end{aligned} \tag{E4.14.13}$$

Infatti, quando $T \to \infty$ la quantità τ/T tende a zero, mentre la funzione $u(T/2 - \tau)$ vale comunque 1 (il punto di applicazione del gradino si sposta verso destra illimitatamente). Geometricamente, si ha la situazione di Figura 4.48. Estendendo per simmetria pari questo risultato, si conclude che, per qualunque τ,

$$R_x(\tau) = \frac{1}{2} \tag{E4.14.14}$$

dalla quale si ricava immediatamente, calcolando la trasformata di Fourier, il risultato (E4.14.9).

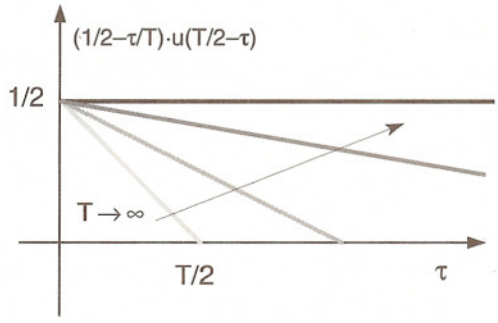


Figura 4.48 Funzione d'autocorrelazione del segnale gradino unitario

❑

4.4.4 Densità spettrale di potenza dei segnali periodici

Una particolare categoria di segnali a potenza finita è rappresentata dai segnali periodici. Per essi vogliamo determinare la funzione di autocorrelazione e la densità spettrale di potenza in funzione del relativo coefficiente di Fourier X_k. Per un segnale $x(t)$ periodico di periodo T_0, l'espressione della funzione di autocorrelazione (4.4.31) si semplifica come segue:

$$R_x(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, x(t-\tau)\, dt \tag{4.4.33}$$

È immediato rendersi conto che la funzione di autocorrelazione di un segnale periodico è periodica in τ dello stesso periodo T_0 del segnale. Per procedere con il calcolo, sostituiamo a $x(t-\tau)$ nella (4.4.33) il relativo sviluppo in serie di Fourier. Si ottiene così

$$R_x(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \sum_{k=-\infty}^{+\infty} X_k\, e^{j2\pi \frac{k}{T_0}(t-\tau)} dt \tag{4.4.34}$$

Invertendo le operazioni di sommatoria e integrazione ricaviamo

$$R_x(\tau) = \sum_{k=-\infty}^{+\infty} X_k \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, e^{j2\pi \frac{k}{T_0}(t-\tau)} dt = \sum_{k=-\infty}^{+\infty} X_k\, e^{-j2\pi \frac{k}{T_0}\tau} \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi\left(-\frac{k}{T_0}\right)t} dt$$

$$= \sum_{k=-\infty}^{+\infty} X_k\, X_{-k}\, e^{-j2\pi k \frac{\tau}{T_0}} \tag{4.4.35}$$

Tenendo conto che il segnale è reale e che quindi $X_{-k} = X_k^*$, si conclude che

$$R_x(\tau) = \sum_{k=-\infty}^{+\infty} X_k \, X_k^* \, e^{-j2\pi k \frac{\tau}{T_0}} = \sum_{k=-\infty}^{+\infty} |X_k|^2 e^{-j2\pi k \frac{\tau}{T_0}} = \sum_{k=-\infty}^{+\infty} |X_k|^2 e^{j2\pi k \frac{\tau}{T_0}} \tag{4.4.36}$$

Dunque, i coefficienti dell'espansione in serie di Fourier della funzione di autocorrelazione del segnale periodico sono dati dal modulo quadro dei coefficienti dell'espansione in serie del segnale stesso:

$$R_x(\tau) \Leftrightarrow |X_k|^2 \tag{4.4.37}$$

Per ricavare la densità spettrale di potenza di $x(t)$ è ora sufficiente calcolare la trasformata continua di Fourier della funzione $R_x(\tau)$. Dalla (4.4.36) si ricava immediatamente

$$S_x(f) = \sum_{k=-\infty}^{+\infty} |X_k|^2 \delta\left(f - \frac{k}{T_0}\right) \tag{4.4.38}$$

Dunque, un segnale periodico ha uno spettro di potenza *a righe* come illustrato in Figura 4.49; ciascuna riga rappresenta il contributo alla potenza del segnale dato dalla componente alla frequenza armonica corrispondente. Pertanto possiamo dire che la potenza, anziché essere *distribuita* con continuità su tutte le frequenze dell'asse reale (come per un generico segnale aperiodico), è *concentrata* sulle frequenze armoniche.

La potenza complessiva del segnale si ottiene integrando la densità spettrale di potenza su tutto l'asse delle frequenze; quindi possiamo scrivere

$$P_x = \int_{-\infty}^{+\infty} S_x(f)\, df = \int_{-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} |X_k|^2 \delta\left(f - \frac{k}{T_0}\right) df = \sum_{k=-\infty}^{+\infty} |X_k|^2 \int_{-\infty}^{\infty} \delta\left(f - \frac{k}{T_0}\right) df = \sum_{k=-\infty}^{\infty} |X_k|^2 \tag{4.4.39}$$

Il risultato espresso dalla (4.4.39), che di seguito riassumiamo, rappresenta la versione del *teorema di Parseval* per i segnali periodici:

$$P_x = \sum_{k=-\infty}^{+\infty} |X_k|^2 \tag{4.4.40}$$

Utilizzando i coefficienti dello sviluppo in serie in forma *reale polare* si ha anche

$$P_x = \sum_{k=-\infty}^{+\infty} |X_k|^2 = \sum_{k=-\infty}^{-1} |X_k|^2 + |X_0|^2 + \sum_{k=1}^{+\infty} |X_k|^2 = A_0^2 + 2\sum_{k=1}^{+\infty} A_k^2 \tag{4.4.41}$$

Osserviamo che la potenza associata alla generica k-esima oscillazione armonica dello sviluppo in forma reale polare è pari a

$$P_k = (2A_k)^2 / 2 = 2A_k^2 \tag{4.4.42}$$

dato che l'oscillazione ha ampiezza di picco $2A_k$. Quindi il teorema di Parseval sancisce che la potenza *totale* del segnale periodico si ottiene semplicemente *sommando* le potenze delle singole oscillazioni armoniche in cui il segnale è scomponibile, come se queste si presentassero *singolarmente*.

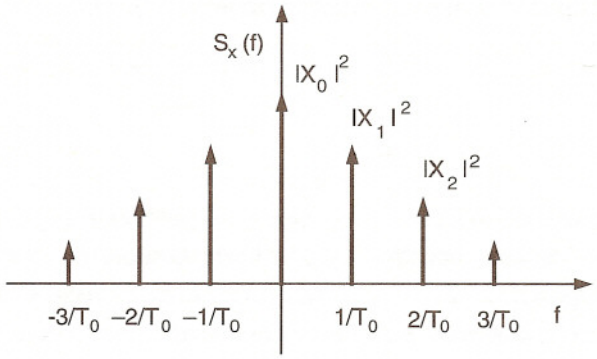


Figura 4.49 Densità spettrale di potenza di un segnale periodico

4.5 Sistemi non lineari

4.5.1 Caratterizzazione dei sistemi non lineari

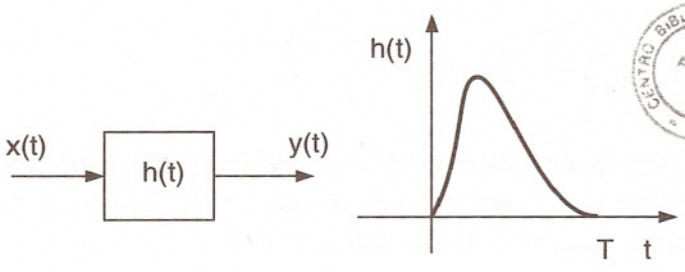
I sistemi lineari stazionari sono caratterizzati dalla conoscenza della risposta impulsiva $h(t)$. Se si considera il particolare SLS causale di Figura 4.50, il segnale $y(t)$ d' uscita è esprimibile come

$$y(t) = \int_{-\infty}^{\infty} x(\alpha)\, h(t-\alpha)\, d\alpha = \int_{t-T}^{t} x(\alpha)\, h(t-\alpha)\, d\alpha \tag{4.5.1}$$

essendo $h(t-\alpha) \neq 0$ solo nell'intervallo $t - T < \alpha < t$. Il valore del segnale d'uscita calcolato all'istante t dipende quindi dai valori che l'ingresso assume in un intervallo di durata T precedente l'istante t stesso. Un comportamento di questo tipo è caratteristico di un *sistema con memoria* e, nel caso specifico, con una memoria *finita* pari a T. Un sistema è invece *senza memoria* se l'uscita a un istante arbitrario dipende solamente dal valore dell'ingresso al medesimo istante.

Pertanto il comportamento ingresso-uscita di un sistema senza memoria è caratterizzato in generale da una relazione del tipo

$$y(t) = \mathcal{T}[x(\alpha), \alpha = t; t] \tag{4.5.2}$$



CENTRO BIBLIOTECARIO INGEGNERIA PISA

Figura 4.50 Sistema lineare stazionario con memoria finita

Se il sistema è anche *stazionario*, l'*operatore* $\mathcal{T}[\cdot]$ "collassa" in una semplice *funzione* $g(x)$ che rappresenta la cosiddetta *caratteristica ingresso-uscita* del sistema (vedi Figura 4.51).

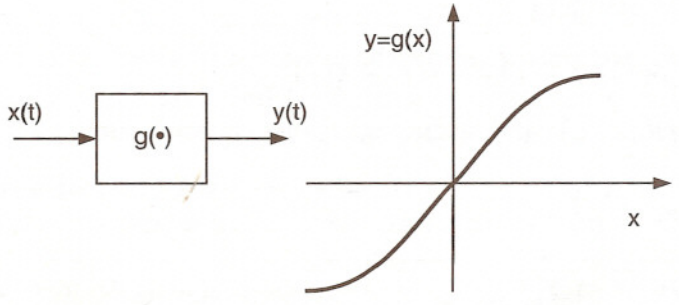


Figura 4.51 Caratteristica ingresso-uscita di un sistema senza memoria

La caratteristica non dipende dal tempo poiché il sistema è stazionario; in tutti gli istanti in cui il segnale di ingresso "passa" da un certo valore x_0, in uscita troveremo comunque il valore $y_0 = g(x_0)$.

Quali sono le peculiarità di funzionamento di un generico sistema non lineare stazionario senza memoria? Quali le principali differenze rispetto a un sistema *lineare*? Suppondendo nota la caratteristica ingresso-uscita $g(x)$, il comportamento della non linearità può essere meglio compreso sviluppando in serie di

Taylor la caratteristica stessa in un intorno dell'origine[2]:

$$g(x)=\sum_{n=0}^{+\infty} g_n\, x^n = g_0 + g_1 x + g_2 x^2 + g_3 x^3 + \cdots \tag{4.5.3}$$

con

$$g_n = \frac{1}{n!}\frac{d^n}{dx^n} g(x)\bigg|_{x=0} \tag{4.5.4}$$

Per semplicità, supponiamo di avere a che fare con una nonlinearità per cui lo sviluppo in serie (4.5.3) arrestato al secondo ordine sia una approssimazione sufficientemente accurata della caratteristica, cioè

$$y(t) = g\big(x(t)\big) = g_0 + g_1 x(t) + g_2 x^2(t) \tag{4.5.5}$$

La presenza del termine quadratico $g_2 x^2(t)$ nell'espressione del segnale di uscita è chiaramente responsabile dell'introduzione di *distorsione non lineare* sul segnale d'ingresso. Per meglio comprendere la natura di tale distorsione, passiamo alle trasformate; la relazione (4.5.5) si traduce nella seguente nel dominio della frequenza:

$$Y(f) = g_0\delta(f) + g_1 X(f) + g_2 X(f) \otimes X(f) \tag{4.5.6}$$

Se, ad esempio, $X(f)$ è uno spettro rettangolare (a banda limitata):

$$X(f) = \mathrm{rect}\left(\frac{f}{2B}\right) \tag{4.5.7}$$

allora la trasformata $Y(f)$ ottenuta calcolando la somma dei tre contributi nella (4.5.6) è quella rappresentata nella Figura 4.52.

Già questo semplice esempio è significativo del meccanismo di distorsione di un sistema nonlineare senza memoria: nel segnale d'uscita ritroviamo, oltre a una componente continua, che in pratica può essere facilmente eliminata con un accoppiamento in alternata del segnale di uscita, anche alcune *componenti frequenziali che non sono presenti nel segnale d'ingresso*, in particolare tutte quelle nella banda $(B, 2B]$. La produzione di queste componenti è chiaramente

[2] Solo caratteristiche con un certo grado di "regolarità" nell'intorno dell'origine possono essere sviluppate in serie di Taylor. Questa condizione non è necessariamente verificata, ma la discussione riveste comunque carattere generale.

dovuta alla presenza del termine *nonlineare* (in particolare quadratico) nella caratteristica ingresso-uscita.

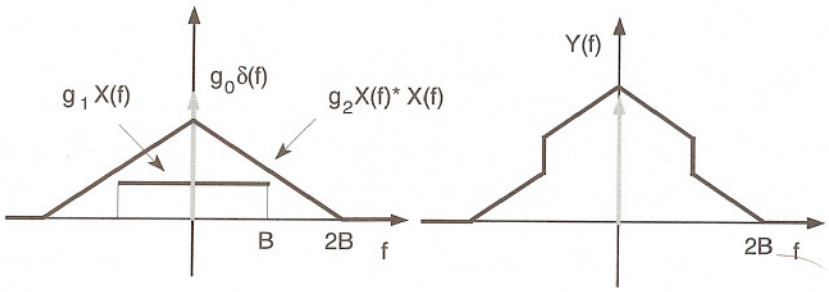


Figura 4.52 Esempio di distorsione non lineare

L'*allargamento dello spettro del segnale*, tipico delle distorsioni non lineari, è tanto più marcato quanto più la caratteristica ingresso-uscita del sistema si discosta da un andamento lineare. Se avessimo preso in considerazione anche un termine *cubico* nell'espansione in serie di Taylor, avremmo ottenuto uno spettro del segnale di uscita con banda *triplicata*, e così via. Osserviamo inoltre che un eventuale filtraggio dell'uscita per reiettare le componenti fuori banda (cioè fuori della banda originaria B) non permette comunque di recuperare una replica fedele del segnale di ingresso; l'effetto della non linearità si manifesta infatti anche all'interno della banda del segnale utile (fenomeno di *intermodulazione*).

4.5.2 Nonlinearità essenziali e parassite

Nella discussione precedente sull'effetto della nonlinearità, essa è stata considerata come un fattore *negativo* producente distorsione sul segnale. Questa visione è in generale valida per sistemi di elaborazione dei segnali, ma deve essere approfondita quando si considerano sistemi nonlineari in generale, in particolare sistemi elettronici di potenza. È importante cioè classificare preliminarmente il tipo di non linearità o, meglio, stabilire se si tratta di una non linearità *essenziale* o *parassita* in relazione alla funzione che il sistema deve svolgere.

Una nonlinearità *parassita* si manifesta nonostante tutti gli sforzi di progettazione in sistemi *nominalmente lineari*, e provoca quindi effetti indesiderati. L'esempio tipico di non linearità parassita è quella presentata da un circuito amplificatore reale. Abbiamo già più volte considerato un amplificatore *ideale*, che è descritto dalla relazione

$$y(t) = A \cdot x(t) \tag{4.5.8}$$

cui corrisponde la caratteristica ingresso-uscita $g(x) = A \cdot x$ rappresentata in Figura 4.53a. Osserviamo esplicitamente che questo sistema rappresenta, per così dire, "l'intersezione" tra la classe dei sistemi lineari stazionari e quella dei sistemi stazionari senza memoria. Per un amplificatore reale, però, la caratteristica non è quella illustrata nella Figura 4.53a, bensì quella (non lineare) di Figura 4.53b, che indica un fenomeno di *saturazione*. Se i valori del segnale d'ingresso sono compresi in un certo intervallo $[-x_M, x_M]$, ove x_M è la *dinamica d'ingresso*, il sistema si comporta come un amplificatore ideale (caratteristica lineare); al di fuori di questo intervallo, il sistema opera in zona non lineare e l'uscita non supera un valore costante y_M (la *dinamica d'uscita*) che rappresenta il livello di saturazione dell'amplificatore. Se si eccede la dinamica di ingresso, si avranno quindi distorsioni nonlineari sul segnale d'uscita.

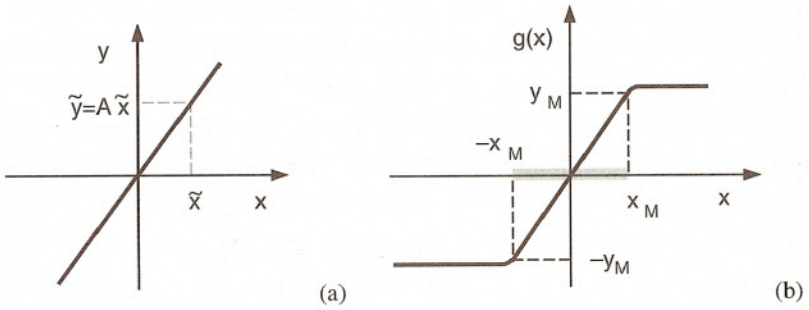


Figura 4.53 Caratteristiche ingresso-uscita dell'amplificatore ideale (a) e reale (b)

Consideriamo al contrario un esempio di nonlinearità *essenziale*. Supponiamo di disporre di un generatore di forma d'onda sinusoidale, e di voler ricavare a partire dal segnale $x(t)$ fornito dal generatore un segnale $y(t)$ costante. Si dispone dunque di un segnale $x(t)$ il cui spettro è costituito da due righe a frequenza $\pm f_0$ e si vuole produrre un segnale $y(t)$ costituito solo da una componente (continua) a frequenza 0. È chiaro che *questa operazione non può essere realizzata con un sistema lineare stazionario.* La relazione fondamentale $Y(f) = H(f)X(f)$ di questi sistemi indica che nello spettro del segnale di uscita non possono esistere componenti che sono di ampiezza nulla nello spettro del segnale di ingresso.

La funzione desiderata, che è quella tipica compiuta dai circuiti *alimentatori* degli apparati elettronici, può essere invece realizzata utilizzando un sistema *non lineare* in cui la nonlinearità diventa parte essenziale. Tale sistema è costituito da due blocchi funzionali ed è rappresentato in Figura 4.54. Il primo blocco è un *raddrizzatore a doppia semionda* con caratteristica ingresso-uscita

$$g(x) = |x| \tag{4.5.9}$$

e il secondo è un filtro passa-basso (quindi un SLS) che estrae dal segnale raddrizzato la sola componente continua.

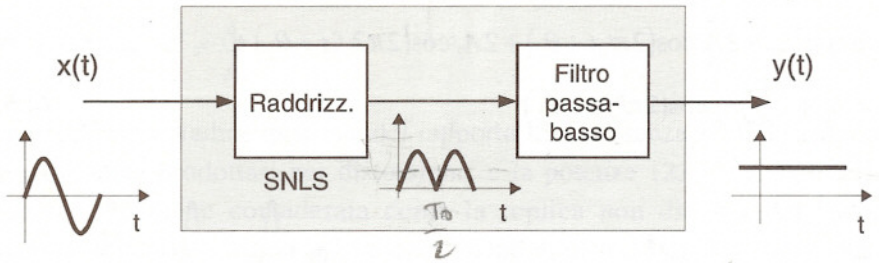


Figura 4.54 Schema a blocchi di un alimentatore

4.5.3 Misura delle distorsioni non lineari

Esistono alcuni metodi di misura per caratterizzare la maggiore o minore influenza della nonlinearità in un sistema in cui questa è vista come parassita. Il più semplice consiste nel dare in ingresso al sistema in esame un segnale sinusoidale

$$x(t) = V_0 \cos(2\pi f_0 t) \tag{4.5.10}$$

Supponiamo dapprima, per semplicità, che la caratteristica ingresso-uscita del sistema sia quella di una nonlinearità del secondo ordine:

$$g(x) = g_0 + g_1 x + g_2 x^2 \tag{4.5.11}$$

Allora il segnale d'uscita è espresso da

$$y(t) = g_0 + g_1 V_0 \cos(2\pi f_0 t) + g_2 V_0^2 \cos^2(2\pi f_0 t) \tag{4.5.12}$$

$$= \left[g_0 + \frac{g_2 V_0^2}{2} \right] + g_1 V_0 \cos(2\pi f_0 t) + \frac{g_2 V_0^2}{2} \cos(2\pi 2 f_0 t) \tag{4.5.13}$$

Nel segnale d'uscita è presente un'oscillazione a frequenza $2f_0$ che non compare nel segnale d'ingresso e che viene chiamata *distorsione di seconda armonica.* Nel caso generale, in cui la caratteristica del sistema è arbitraria, in risposta al segnale sinusoidale periodico di periodo $T_0 = 1/f_0$, il sistema produrrà il segnale

$$y(t) = g[x(t)] = g[x(t + T_0)] = y(t + T_0) \tag{4.5.14}$$

Il segnale di uscita, visto che il sistema è senza memoria, è a sua volta periodico con lo stesso periodo di quello d'ingresso, come esemplificato in Figura 4.55, anche se non sinusoidale. Ciò consente di sviluppare $y(t)$ in serie di Fourier (forma reale polare):

$$y(t) = A_0 + 2A_1 \cos(2\pi f_0 t + \theta_1) + 2A_2 \cos(2\pi 2 f_0 t + \theta_2) + \cdots$$
$$= A_0 + 2\sum_{k=1}^{+\infty} A_k \cos(2\pi k f_0 t + \theta_k) \qquad (4.5.15)$$

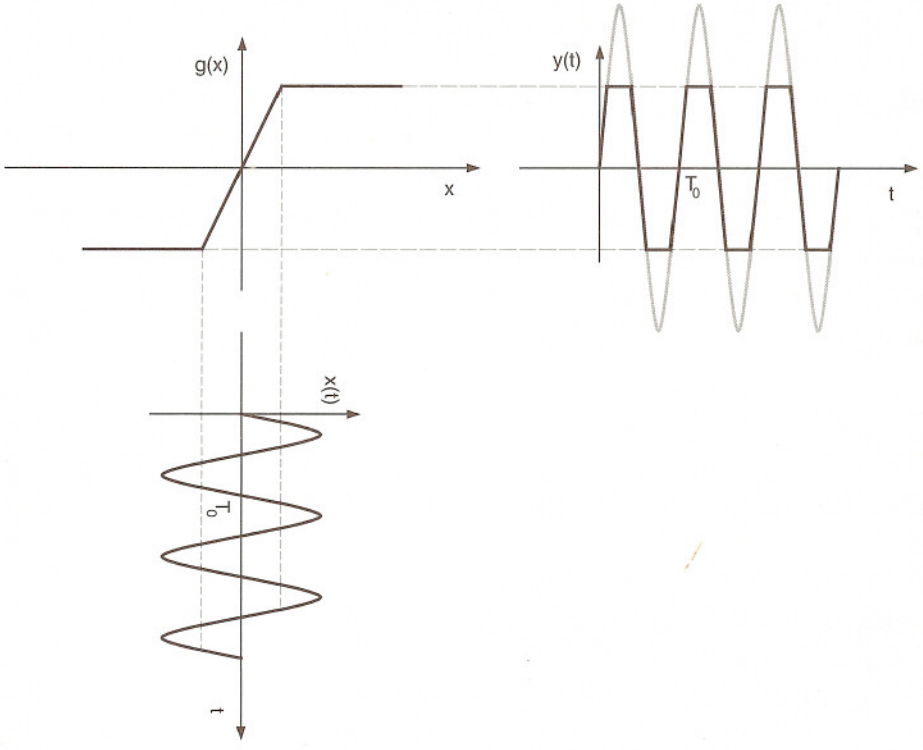


Figura 4.55 Esempio di distorsione nonlineare (amplificatore con saturazione)

Come già detto, l' effetto della non linearità è quello di *generare componenti non presenti nel segnale d'ingresso*; le componenti generate sono qui in *relazione armonica* con la frequenza del segnale d'ingresso; questo meccanismo di distorsione si indica dunque con il nome di *distorsione armonica*. La componente continua A_0 non viene considerata come prodotto di distorsione (è spesso ininfluente nelle applicazioni), e la componente per $k = 1$ è in pratica la parte

"utile" del segnale di uscita perché replica fedele del segnale di ingresso. Per quantificare l'influenza della nonlinearità si definisce allora il *coefficiente di distorsione di k-esima armonica*

$$D_k \triangleq \sqrt{\frac{(2A_k)^2/2}{(2A_1)^2/2}} = \frac{A_k}{A_1} \quad , \quad k \geq 2 \tag{4.5.16}$$

che rappresenta la radice quadrata del rapporto tra la potenza $(2A_k)^2/2$ della k-esima armonica prodottasi per distorsione, e la potenza $(2A_1)^2/2$ della prima armonica, che viene considerata come la replica non distorta del segnale d'ingresso. In teoria esistono infiniti coefficienti di distorsione, benché in pratica siano rilevanti soltanto quelli di ordine inferiore (seconda e terza armonica).

Per disporre comunque di un'informazione sulla distorsione complessiva, cioè relativa a *tutte* le armoniche prodotte, si definisce innanzitutto il *segnale distorsione* come "residuo" del segnale di uscita, tolta la componente utile e quella (ininfluente) continua:

$$d(t) = y(t) - A_0 - 2A_1 \cos(2\pi f_0 t + \theta_1) = 2\sum_{k=2}^{+\infty} A_k \cos(2\pi k f_0 t + \theta_k) \tag{4.5.17}$$

Si definisce poi il *coefficiente di distorsione armonica totale* (THD, *Total Harmonic Distortion*) come il rapporto tra la potenza del residuo di distorsione $d(t)$ e quella della componente utile:

$$D \triangleq \sqrt{\frac{P_d}{(2A_1)^2/2}} = \frac{d_{eff}}{\sqrt{2}A_1} \tag{4.5.18}$$

Si noti che, per il teorema di Parseval:

$$P_d = 2\sum_{k=2}^{+\infty} A_k^2 \quad , \quad P_y = A_0^2 + 2\sum_{k=1}^{+\infty} A_k^2 \Rightarrow P_d = P_y - A_0^2 - 2A_1^2 \tag{4.5.19}$$

Allora il coefficiente di distorsione totale (4.5.18) risulta

$$D = \sqrt{\frac{P_y - A_0^2 - 2A_1^2}{2A_1^2}} = \sqrt{\frac{P_y - A_0^2}{2A_1^2} - 1} \tag{4.5.20}$$

Esempio 4.15

Il segnale $x(t) = \cos(2\pi f_0 t)$ è applicato al sistema nonlineare di Figura 4.56.

Determiniamo i coefficienti di distorsione di seconda armonica D_2, di terza armonica D_3 e di distorsione totale D del segnale di uscita $y(t)$. I segnali di ingresso e di uscita del sistema in esame sono rappresentati in Figura 4.57.

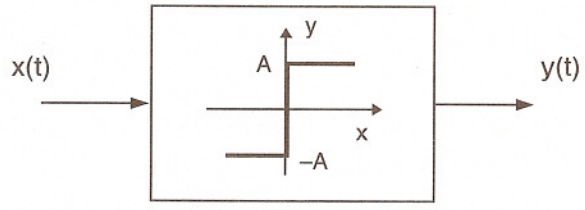


Figura 4.56 Limitatore/Comparatore

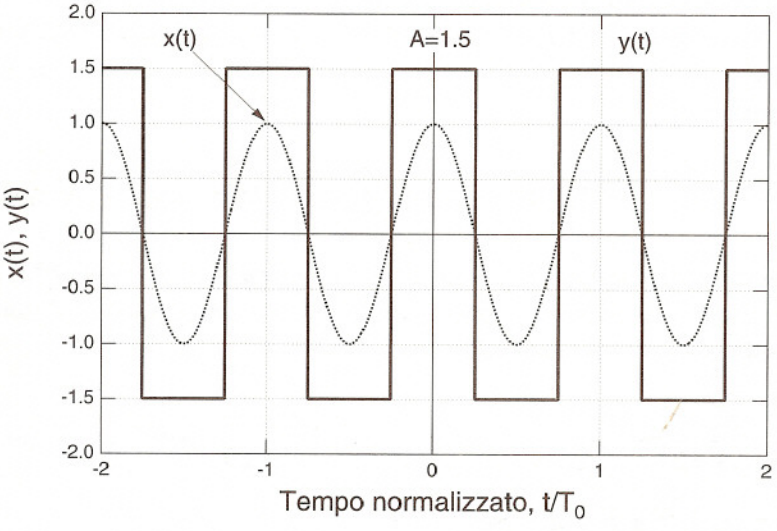


Figura 4.57 Andamento dei segnali di ingresso e di uscita del sistema di Figura 4.56

Il segnale di uscita è un'onda quadra di periodo $T_0 = 1/f_0$ e di ampiezza pari ad A. Come dimostrato nell'Esempio 2.4, i coefficienti dello sviluppo in serie di Fourier del segnale $y(t)$ sono

$$Y_k = \begin{cases} 0 & \text{per } k = 0 \\ A\,\text{sinc}(k/2) & \text{altrimenti} \end{cases} \qquad \text{(E4.15.1)}$$

Poiché $Y_2 = 0$ (il segnale $y(t)$ è alternativo) il coefficiente di distorsione di seconda armonica è nullo. Il coefficiente di distorsione di terza armonica è invece dato da

$$D_3 \stackrel{\Delta}{=} \frac{A_3}{A_1} = \frac{|Y_3|}{|Y_1|} = \frac{A\,|\mathrm{sinc}(3/2)|}{A\,|\mathrm{sinc}(1/2)|} = \frac{1}{3} \qquad \text{(E4.15.2)}$$

Inoltre, la potenza del segnale di uscita è banalmente $P_y = A^2$ e il valore medio del segnale è nullo ($Y_0 = 0$); la potenza del residuo di distorsione P_d è allora pari a:

$$P_d = P_y - 2|Y_1|^2 = A^2 - 2\,A^2\,\mathrm{sinc}^2(1/2) = A^2\left[1 - \frac{8}{\pi^2}\right] \qquad \text{(E4.15.3)}$$

Allora il coefficiente di distorsione armonica totale è dato da

$$D = \sqrt{\frac{P_d}{2|Y_1|^2}} = \sqrt{\frac{A^2\left[1-8/\pi^2\right]}{2A^2\,\mathrm{sinc}^2(1/2)}} = \sqrt{\frac{1-8/\pi^2}{8/\pi^2}} = \sqrt{\frac{\pi^2}{8} - 1} \cong 0.483 \qquad \text{(E4.15.4)}$$

❑

Esempio 4.16

Il segnale $x(t) = 2A\cos(2\pi f_0 t)$ è applicato al sistema nonlineare di Figura 4.58. Calcoliamo i coefficienti di distorsione di seconda armonica D_2, di terza armonica D_3 e di distorsione armonica totale D del segnale di uscita $y(t)$.

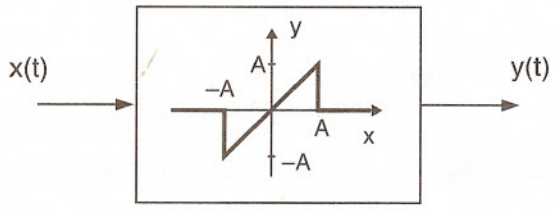


Figura 4.58 Sistema nonlineare

Il segnale di ingresso e il corrispondente segnale di uscita sono rappresentati in Figura 4.59. Si vede che il segnale di uscita è alternativo, quindi tutte le sue armoniche di ordine pari hanno ampiezza nulla e, come nell'esempio precedente, il coefficiente di distorsione di seconda armonica D_2 è nullo. Per calcolare il coefficiente di distorsione di terza armonica D_3 è necessario determinare i coefficienti Y_3 e Y_1. Essendo $y(t)$ un segnale pari, il k-esimo coefficiente del suo sviluppo in serie di Fourier può essere calcolato con la formula semplificata (2.6.9):

$$Y_k = \frac{2}{T_0}\int_0^{T_0/2} y(t)\cos(2\pi k f_0 t)\,dt \qquad \text{(E4.16.1)}$$

da cui segue

$$Y_1 = \frac{2}{T_0}\int_0^{T_0/2} y(t)\cos(2\pi f_0 t)\,dt = \frac{2}{T_0}\int_{T_0/6}^{T_0/3} y(t)\cos(2\pi f_0 t)\,dt$$
$$= \frac{4A}{T_0}\int_{T_0/6}^{T_0/3} \cos^2(2\pi f_0 t)\,dt = A\left(\frac{1}{3} - \frac{\sqrt{3}}{2\pi}\right) \qquad \text{(E4.16.2)}$$

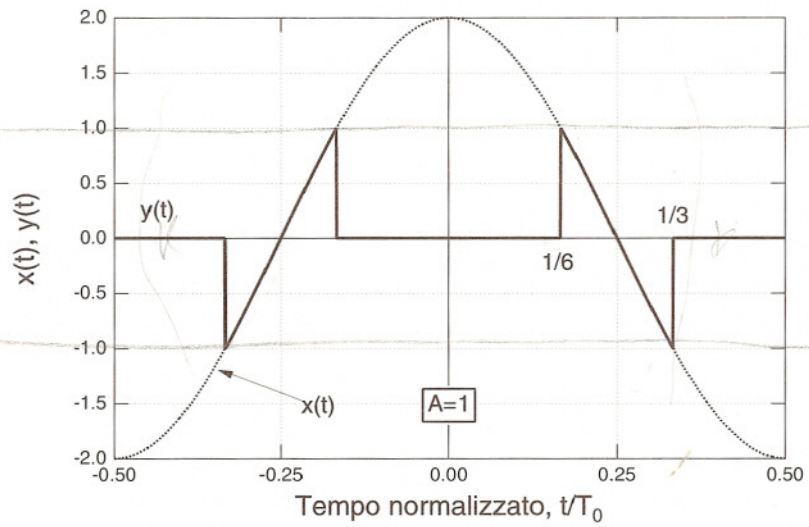


Figura 4.59 Andamento dei segnali di ingresso e di uscita del sistema di Figura 4.58

Analogamente,

$$Y_3 = \frac{2}{T_0}\int_0^{T_0/2} y(t)\cos(6\pi f_0 t)\,dt = \frac{2}{T_0}\int_{T_0/6}^{T_0/3} y(t)\cos(6\pi f_0 t)\,dt$$
$$= \frac{4A}{T_0}\int_{T_0/6}^{T_0/3} \cos(2\pi f_0 t)\cos(6\pi f_0 t)\,dt$$
$$= \frac{4A}{T_0}\int_{T_0/6}^{T_0/3} \frac{1}{2}\left[\cos(4\pi f_0 t) + \cos(8\pi f_0 t)\right]dt = -\frac{A\sqrt{3}}{4\pi} \qquad \text{(E4.16.3)}$$

Usando i risultati (E4.16.2-3) si ottiene

$$D_3 = \frac{3\sqrt{3}}{2\left[2\pi - 3\sqrt{3}\right]} \cong 2.39 \qquad \text{(E4.16.4)}$$

Procediamo ora con il calcolo del coefficiente di distorsione totale. Essendo $y(t)$ un segnale pari, la sua potenza media è data da

$$P_y = \frac{2}{T_0}\int_0^{T_0/2} y^2(t)\,dt = \frac{2}{T_0}\int_{T_0/6}^{T_0/3} y^2(t)\,dt$$
$$= \frac{8A^2}{T_0}\int_{T_0/6}^{T_0/3} \cos^2\left(2\pi f_0 t\right)dt = 2A^2\left(\frac{1}{3} - \frac{\sqrt{3}}{2\pi}\right) \qquad \text{(E4.16.5)}$$

La potenza associata alle componenti di distorsione è invece

$$P_d = P_y - |Y_0|^2 - 2|Y_1|^2 = 2A^2\left(\frac{1}{3} - \frac{\sqrt{3}}{2\pi}\right) - 2A^2\left(\frac{1}{3} - \frac{\sqrt{3}}{2\pi}\right)^2 \qquad \text{(E4.16.6)}$$

($Y_0 = 0$). Allora il coefficiente di distorsione totale risulta

$$D = \sqrt{\frac{P_d}{2|Y_1|^2}} = \sqrt{\frac{6\pi}{2\pi - 3\sqrt{3}} - 1} \cong 4.16 \qquad \text{(E4.16.7)}$$

❑

Sommario

Un sistema monodimensionale (a un ingresso e una uscita) *trasforma* un segnale d'ingresso $x(t)$ (sollecitazione) in un segnale d'uscita (risposta) $y(t)$. Anche senza conoscere la *struttura* del sistema, è possibile individuare alcune proprietà di quest'ultimo: *stazionarietà* se il comportamento del sistema non varia nel tempo; *causalità* se l'evoluzione dell'uscita non dipende dalla futura evoluzione dell'ingresso; *istantaneità* se il valore dell'uscita dipende solo dal valore dell'ingresso al medesimo istante; *stabilità* se a qualunque ingresso limitato in ampiezza corrisponde sempre un'uscita limitata in ampiezza; *invertibilità* se il segnale d'ingresso può sempre essere ricostruito a partire dall'osservazione del segnale d'uscita; *linearità* se al sistema è applicabile il principio di sovrapposizione degli effetti. Limitandoci allo studio dei sistemi lineari stazionari (SLS), si definisce la *risposta impulsiva* $h(t)$ come risposta all'eccitazione $\delta(t)$, e si di-

mostra che questa caratterizza completamente il comportamento del sistema; in particolare, l'uscita del sistema si può calcolare come $y(t) = x(t) \otimes h(t)$. Una caratterizzazione equivalente si ottiene in ambito frequenziale definendo la *risposta in frequenza* del sistema $H(f)$, cioè la trasformata di Fourier della risposta impulsiva; la risposta in frequenza può anche essere calcolata come rapporto tra le trasformate dei segnali d'uscita e d'ingresso $H(f) = Y(f)/X(f)$, oppure come il coefficiente (complesso) secondo cui viene modificata un'oscillazione complessa alla frequenza f che passa attraverso il sistema: $y(t) = H(f)e^{j2\pi ft}$.

La caratterizzazione dei SLS in ambito frequenziale risulta indispensabile quando il sistema stesso deve effettuare operazioni di elaborazione dei segnali *selettive in frequenza*, cioè operazioni di *filtraggio*. A questo proposito si definiscono le risposte dei filtri ideali passa-basso, passa-alto, passa-banda ed elimina-banda. Questi sistemi fondamentali sono caratterizzati da una banda *passante* e una banda *oscura* che ne rappresentano le caratteristiche fondamentali. Si dimostra però che questi filtri *ideali* non sono fisicamente realizzabili perché *non causali*, come si può verificare attraverso il criterio di Paley-Wiener sulla risposta in ampiezza.

Il concetto di banda introdotto per i filtri può essere generalizzato e precisato anche per segnali arbitrari a energia finita. Attraverso opportune definizioni di *banda* e *durata* si arriva al risultato che queste due quantità variano in maniera inversa. In particolare, non è possibile ridurre arbitrariamente la banda di un segnale senza aumentarne in proporzione la durata: il prodotto di queste due grandezze è *limitato inferiormente*. Le approssimazioni dei filtri ideali che si realizzano in pratica possono però introdurre *distorsioni lineari* sul segnale utile. Il criterio di non distorsione richiede che la risposta in ampiezza del filtro sia costante (piatta) e la risposta in fase sia proporzionale alla frequenza (lineare) *all'interno della banda del segnale utile*.

Attraverso lo studio dei SLS è possibile anche caratterizzare il contenuto energetico dei segnali in ambito frequenziale. Il teorema di Parseval, nelle varie versioni per segnali periodici e aperiodici, permette di calcolare l'energia (potenza) dei segnali a partire dalla conoscenza dello spettro di ampiezza. Le funzioni *densità spettrale di energia* e di *potenza* (rispettivamente per i segnali a energia e potenza finita) rappresentano, frequenza per frequenza, il contributo "locale" all'energia (potenza) del segnale fornito dalle componenti in un intorno della frequenza considerata, rapportato all'ampiezza dell'intorno stesso. Quando il segnale viene filtrato, queste funzioni vengono modificate in ragione della *risposta in ampiezza al quadrato* del sistema, $|H(f)|^2$. Il teorema di Wiener-

Khintchine permette poi di definire le densità spettrali come *trasformata di Fourier* delle rispettive funzioni di autocorrelazione.

I sistemi nonlineari senza memoria sono caratterizzabili attraverso la *caratteristica ingresso-uscita* $g(x)$, e si distinguono dai SLS perché in generale producono componenti spettrali non presenti nel segnale d'ingresso. La maniera più semplice per quantificare in questi sistemi l'influenza delle *distorsioni non lineari* (considerate come fenomeno parassita) è la misura dei *coefficienti di distorsione armonica* quando l'ingresso del sistema è sinusoidale.

Esercizi proposti

4.1 Per ciascuno dei seguenti segnali:

$$x_1(t) = 2 + \cos(7\pi t/T)$$
$$x_2(t) = \text{sen}(5\pi t/T) + 3\cos(7\pi t/T)$$
$$x_3(t) = 2\cos(\pi t/T) - \text{sen}(3\pi t/T)$$

dire (giustificando il risultato) se ed eventualmente in che modo essi vengono distorti nel passaggio attraverso il sistema lineare stazionario la cui risposta in frequenza è rappresentata in Figura 4.60.

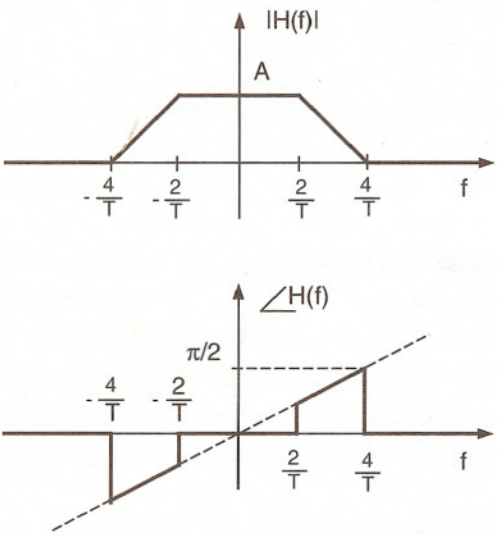


Figura 4.60

4.2 Calcolare la funzione di autocorrelazione dei due segnali $x(t)$ e $y(t)$ in Figura 4.44.

4.3 Partendo dalla relazione ($x(t)$ segnale a energia finita)

$$\int_{-\infty}^{\infty}[x(t) \pm x(t-\tau)]^2 dt \geq 0$$

dimostrare la proprietà (4.4.25) $| R_x(\tau) | \leq R_x(0)$.

4.4 Il teorema di Parseval può anche essere generalizzato considerando due segnali $x(t)$ e $y(t)$ a energia finita. Se con $X(f)$ e $Y(f)$ si indicano le rispettive trasformate di Fourier di questi segnali, dimostrare che (*teorema di Parseval generalizzato)*

$$\int_{-\infty}^{\infty} x(t)\, y^*(t) dt = \int_{-\infty}^{\infty} X(f)\, Y^*(f)\, df$$

4.5 Considerando lo schema di Figura 4.61, in cui il segnale $x(t)$ è espresso da

$$x(t) = \text{sinc}\left(\frac{t - T/2}{T}\right) + \text{sinc}\left(\frac{t + T/2}{T}\right)$$

determinare l'espressione del segnale di uscita $y(t)$ sapendo che

$$c(t) = \sum_{k=-\infty}^{+\infty} \text{rect}\left(\frac{t - kT}{T/2}\right)$$

e che le risposte in frequenza $H_1(f)$ e $H_2(f)$ sono come in Figura 4.62.

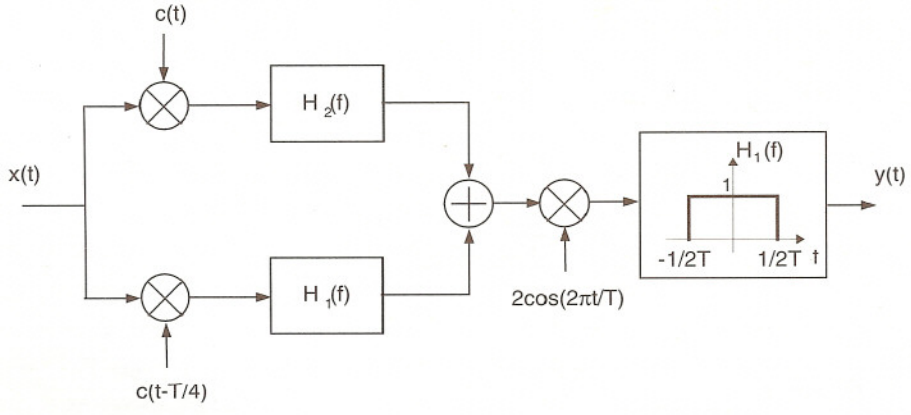


Figura 4.61

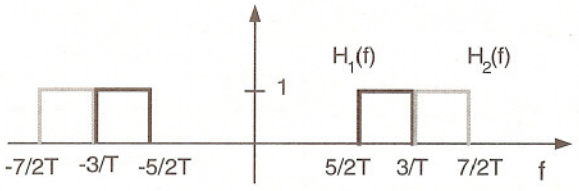


Figura 4.62

4.6 Il segnale

$$y(t) = \sum_{k=-\infty}^{+\infty} (-1)^k x(t - kT)$$

con

$$x(t) = \exp[-t/T]\,\mathrm{rect}\left(\frac{t - T/2}{T}\right)$$

viene applicato in ingresso al sistema di Figura 4.63 (la caratteristica della nonlinearità è $w = y^2$). Determinare la potenza P_z del segnale di uscita $z(t)$.

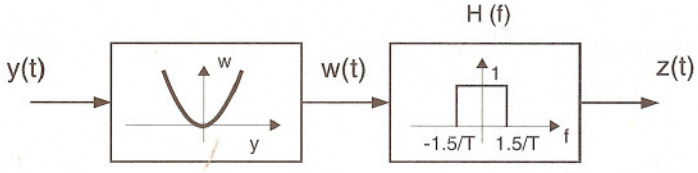


Figura 4.63

4.7 Calcolare l'energia E_x e la potenza P_x del segnale aperiodico $x(t) = 2\,\mathrm{sinc}^2(Bt)\sin(2\pi Bt)$.

4.8 Determinare direttamente la densità spettrale di potenza $S_x(f)$ del segnale $x(t) = \cos(2\pi f_0 t)$ senza sfruttarne la proprietà di periodicità.

4.9 Determinare la caratteristica ingresso-uscita di un sistema senza memoria *lineare e stazionario*. È possibile in questo caso calcolarne la risposta impulsiva?

4.10 Il segnale $x(t) = 2\cos(2\pi Bt)$ è applicato al sistema di Figura 4.64. Sapendo che $h(t) = 2B\,\mathrm{sinc}^2(2Bt)$, determinare la potenza P_z del segnale di uscita $z(t)$.

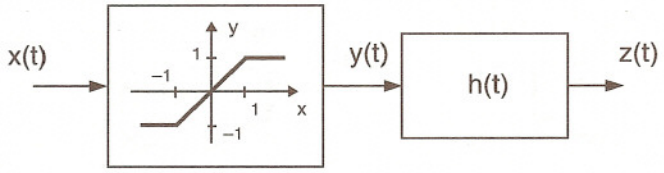


Figura 4.64

4.11 Il segnale $x(t)=\sqrt{2}a\cos(2\pi f_0 t)$ è applicato al sistema nonlineare di Figura 4.65. Calcolare il coefficiente di distorsione armonica totale D del segnale di uscita $y(t)$.

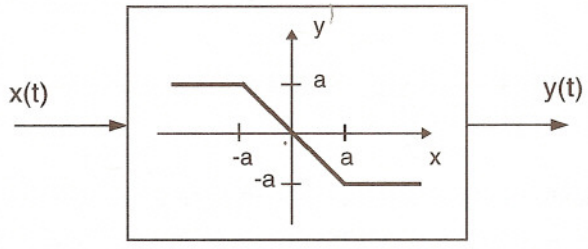


Figura 4.65

4.12 La funzione di *correlazione incrociata* tra due segnali a energia finita $x(t)$ ed $y(t)$ è definita come segue:

$$R_{xy}(\tau)\triangleq\int_{-\infty}^{+\infty}x(t)y(t-\tau)dt$$

Dimostrare che questa funzione, introdotta per misurare la rassomiglianza tra due segnali al variare del ritardo τ, gode della proprietà

$$R_{xy}(\tau)=-R_{yx}(\tau)$$

A che cosa è uguale questa funzione quando $x(t)\equiv y(t)$?

4.13 Dimostrare il teorema di Parseval per segnali periodici (4.4.38) direttamente senza usare l'espressione (4.4.40) della densità spettrale di potenza. [*Suggerimento: si parta dall'espressione della potenza per un segnale periodico, si scriva che* $x^2(t)=x(t)\cdot x^*(t)$ *e poi...*]

4.14 Trovare tramite il teorema di Wiener-Khintchine la densità spettrale di potenza $S_x(f)$ del segnale $x(t)=\mathrm{sgn}(t)$.

4.15 Calcolare la banda a –3 dB B_{-3} del sistema lineare stazionario rappresentato in Figura 4.66 ($\omega_0 = 2\pi f_0$).

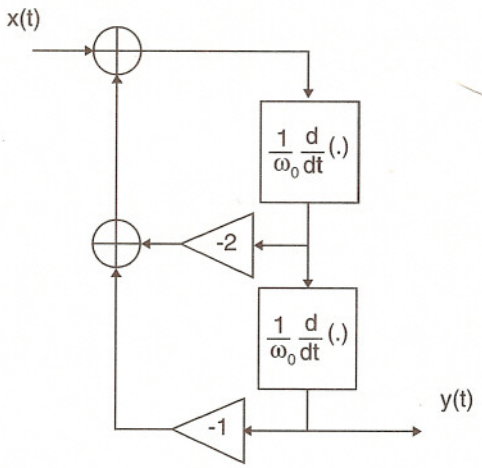


Figura 4.66

4.16 Determinare e rappresentare la risposta $g(t)$ al gradino unitario u(t) per il SLS di Figura 4.67.

4.17 Un sistema lineare stazionario è caratterizzato dalla risposta in frequenza

$$H(f) = \exp\left(-\frac{\ln 2}{8 f_0^2} f^2\right)$$

Calcolare la banda a –3 dB B_{-3} di tale sistema e determinarne poi la risposta all'eccitazione $x(t) = \text{rect}(2 f_0 t)$. [*Suggerimento: usare la funzione* $\Phi(\cdot)$ *definita per una densità di probabilità Gaussiana standard*]

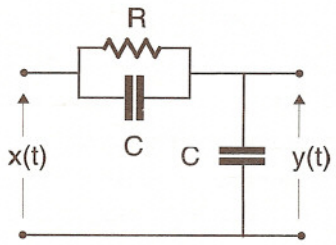


Figura 4.67

5

Segnali a tempo discreto

FINESTRA MOBILE

5.1 Dal tempo continuo al tempo discreto

5.1.1 Campionamento dei segnali a tempo continuo

Nel Capitolo 1 è stata discussa la classificazione dei segnali in base al tipo della variabile indipendente, che in genere è identificata con una grandezza di carattere temporale. Sappiamo già che un *segnale a tempo discreto* è una successione x_n o *sequenza* $x[n]$ di numeri, ed è quindi rappresentabile con una funzione di variabile intera relativa avente valori reali o complessi.

Supponiamo di compiere alcune osservazioni di traffico automobilistico autostradale: a un casello di uscita, annotiamo l'orario di ingresso di ogni vettura, misurato in secondi a partire dalle ore 0.00, e riportiamo questi dati in una tabella (Tabella 5.1). Essa è composta dal numero d'ordine n dell'automobile in ingresso al casello, e dal relativo dato orario $x[n]$: la tabella rappresenta un segnale a tempo discreto, che può essere elaborato per ricavare informazioni sul progetto e il dimensionamento del sistema di riscossione dei pedaggi.

Un caso più tipico nell'elaborazione dei segnali è quello in cui il segnale a tempo discreto viene ottenuto da un segnale a tempo continuo attraverso la cosiddetta operazione di *campionamento*. *Campionare* un segnale $x(t)$ significa "estrarre" dal segnale stesso i valori che esso assume a istanti temporali equispaziati, cioè multipli di un intervallo T detto *periodo di campionamento*, come viene illustrato in Figura 5.1. Con questa operazione viene a crearsi una *sequenza* il cui valore n-esimo $x[n]$ è il valore assunto dal segnale a tempo continuo all'istante nT:

$$x[n] = x(nT) \tag{5.1.1}$$

Tabella 5.1 Esempio di sequenza

n	$x[n]$
0	0
1	15.2
2	33.9
3	65.4
4	68.2
5	129.4
6	162.4
7	312.9
8	423.5
9	428.5
10	629.2
11	734.7
⋮	⋮

Nella Figura 5.1, l'operazione di campionamento viene simbolicamente effettuata da un dispositivo, il *campionatore*, indicato con una sorta di "interruttore" che si chiude per un intervallo di durata infinitesima. La *cadenza* con cui l' interruttore si chiude, cioè con la quale il segnale viene campionato, è pari a

$$f_c = \frac{1}{T} \tag{5.1.2}$$

e prende il nome di *frequenza di campionamento* (*sampling frequency*), misurata in Hz o in campioni/s.

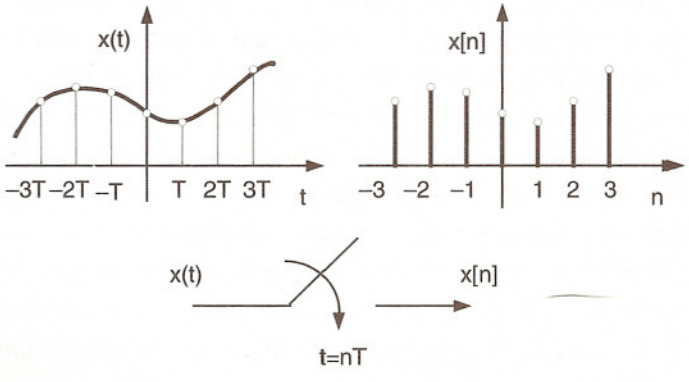


Figura 5.1 Campionamento di un segnale analogico

Nella pratica, come già menzionato nel Capitolo 1, l'operazione di campionamento viene effettuata dai *convertitori analogico/digitale* (comunemente detti *convertitori A/D*). Questi dispositivi sono comandati da un segnale di *clock* (temporizzazione) alla frequenza f_c, che fornisce gli impulsi di comando al circuito per effettuare le varie operazioni di campionamento. Il campionatore *ideale* di Figura 5.1 estrae in corrispondenza di ogni impulso di clock il valore del segnale di ingresso $x(t)$ all'istante di campionamento, che è in generale un numero reale con infinite cifre decimali. Diversamente dal campionatore ideale, il convertitore A/D rende invece una *rappresentazione finita* di questo numero reale (segnale *numerico*), e precisamente in aritmetica binaria su un numero finito di cifre (bit), variabile in genere da 8 a 16. In tutto questo capitolo non ci occuperemo più del piccolo errore insito nella rappresentazione del numero reale su un numero finito di cifre, detto *errore di quantizzazione*, ma supporremo sempre di effettuare operazioni di campionamento *ideali*.

Lo studio e l'elaborazione dei segnali a tempo discreto ha assunto grande importanza in questi ultimi per la possibilità di utilizzare componenti numerici ad alta velocità e affidabilità, e a basso costo. La tendenza moderna, risultante dai grandi progressi compiuti dai componenti elettronici ad altissima integrazione (VLSI, *Very Large-Scale Integration*), è quella di usare per l'elaborazione dei segnali a tempo discreto dei *microprocessori specializzati* (dedicati) chiamati DSP (*Digital Signal Processor*, elaboratore numerico di segnali). L'elaborazione del segnale viene eseguita sui valori *digitali* estratti dal segnale stesso tramite conversione A/D, e si risolve nell'esecuzione di un opportuno *programma* da parte del microprocessore. Questa struttura, rappresentata in Figura 5.2, è estremamente flessibile, nel senso che diverse funzioni di elaborazione possono essere realizzate semplicemente cambiando il *programma* di elaborazione (software) senza dover minimamente modificare la struttura fisica (hardware) del circuito.

Figura 5.2 Elaborazione numerica di un segnale a tempo continuo

Il segnale a tempo discreto $y[n]$ risultante dall'elaborazione numerica deve poi essere *riconvertito* in forma analogica (cioè in un segnale a tempo continuo). Questa operazione è la *conversione digitale/analogico* (D/A) indicata in Figura

5.2. Dal punto di vista teorico, il dispositivo che produce in uscita un segnale a tempo continuo a partire da un segnale d'ingresso a tempo discreto è chiamato *interpolatore*, e il suo funzionamento è schematizzato in Figura 5.3. Le operazione *duali* di campionamento e interpolazione verranno analizzate più approfonditamente nei Paragrafi 5.3 e 5.4.

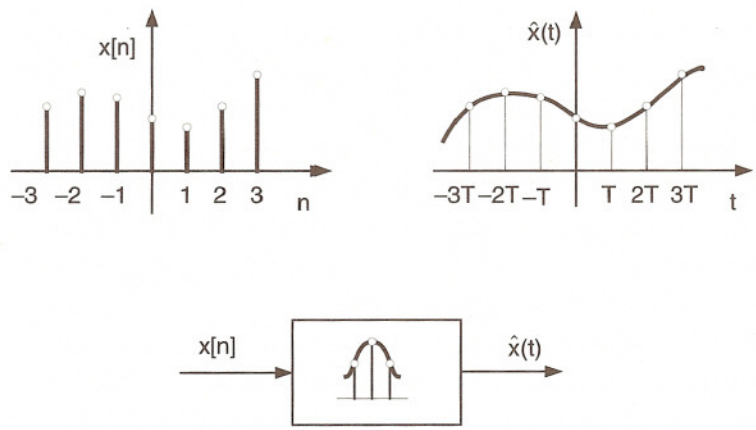


Figura 5.3 Interpolazione di un segnale a tempo discreto

5.1.2 Alcuni segnali notevoli

Consideriamo adesso qualche esempio di segnali a tempo discreto (brevemente: segnali discreti) di particolare rilevanza e utilità. Alcuni di essi possono essere derivati dalle rispettive controparti a tempo continuo per semplice campionamento, altri necessitano invece di importanti precisazioni.

i) sequenza *gradino unitario* (Figura 5.4):

$$u[n] = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases} \tag{5.1.3}$$

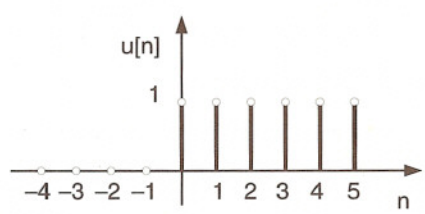


Figura 5.4 Sequenza gradino unitario

Per ovvi motivi, non si pone per la sequenza $u[n]$ la questione del valore nel punto di discontinuità presentata invece dal gradino a tempo continuo $u(t)$.
ii) sequenza *esponenziale unilatera* (Figura 5.5):

$$x[n] = a^n u[n] \tag{5.1.4}$$

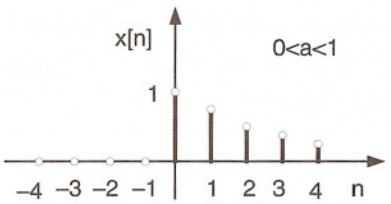


Figura 5.5 Sequenza esponenziale unilatera

iii) sequenza δ *o impulsiva* (Figura 5.6):

$$\delta[n] = \begin{cases} 1 & n = 0 \\ 0 & \text{altrimenti} \end{cases} \tag{5.1.5}$$

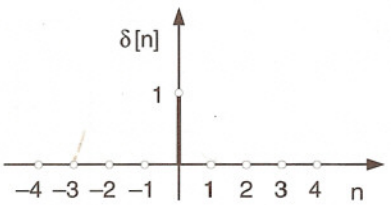


Figura 5.6 Sequenza impulsiva

La sequenza $\delta[n]$ è in pratica ciò che in matematica è chiamato "simbolo di Kronecker" $\delta_{n,k}$:

$$\delta_{n,k} = \delta[n-k] = \begin{cases} 1 & n = k \\ 0 & \text{altrimenti} \end{cases} \tag{5.1.6}$$

Non ha senso pensare di ottenere la sequenza δ per campionamento del segnale analogico $\delta(t)$. Si ricordi che il valore di quest'ultimo per $t = 0$ non è limitato, e quindi sarebbe *impossibile* (o meglio, privo di senso) definire $\delta[0] = \delta(0)$.

Le sequenze $u[n]$ e $\delta[n]$ sono legate da relazioni simili a quelle che legano $u(t)$ e $\delta(t)$. Infatti, come il segnale gradino è la funzione integrale della $\delta(t)$ (si

veda la (3.4.9)), così nell'ambito del tempo discreto si ha

$$u[n] = \sum_{i=-\infty}^{n} \delta[i] \tag{5.1.7}$$

cioè il segnale gradino discreto è la *sequenza somma* della sequenza $\delta[n]$. Dualmente, sappiamo che la funzione $\delta(t)$ è la derivata temporale del segnale $u(t)$. Nell'ambito del tempo discreto non è ovviamente definita l'operazione di derivazione, e la relazione analoga coinvolge la cosiddetta "differenza all'indietro del primo ordine" o *incremento*:

$$\delta[n] = u[n] - u[n-1] \tag{5.1.8}$$

iv) sequenza *impulso rettangolare causale di durata N* (Figura 5.7a):

$$x[n] = u[n] - u[n-N] \tag{5.1.9}$$

L'impulso rettangolare è qui ottenuto (si ricordi l'Esempio 4.8) come *differenza* tra la sequenza $u[n]$ e la stessa sequenza ritardata di N campioni (Figura 5.7.b).

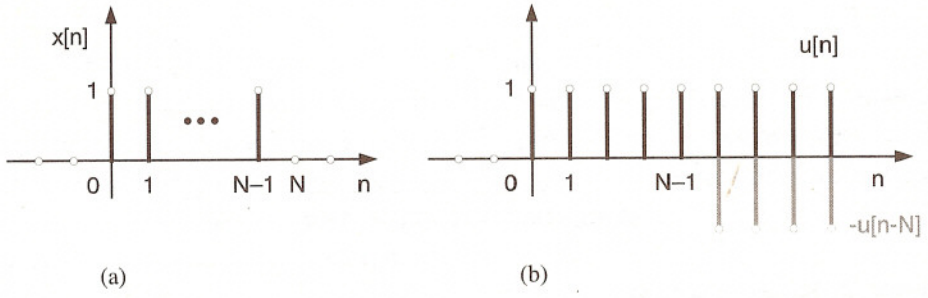


Figura 5.7 Impulso rettangolare discreto (a) e sua costruzione (b)

v) *oscillazione complessa discreta alla frequenza normalizzata* F_0:

$$x[n] = \exp(j2\pi F_0 n) \tag{5.1.10}$$

che è la versione a tempo discreto della oscillazione complessa analogica $\exp(j2\pi f_0 t)$ alla frequenza f_0. La frequenza F_0 è *normalizzata* perché rappresenta un numero puro. Si noti che, rispetto alla variabile temporale n, questo segnale è periodico *solo se la frequenza* F_0 *è un numero razionale,* cioè se $F_0 = p/q$, p e q interi. Il lettore trovi il periodo N_0 dell'oscillazione $x[n]$ sotto quest'ultima ipotesi.

delle frequenze normalizzate di ampiezza unitaria, ad esempio per $F \in [-1/2, 1/2]$ che chiameremo *intervallo-base*. Inoltre, la definizione (5.2.1) di trasformata è più che esauriente ai fini matematici (ed è infatti adottata in molti testi di teoria dei segnali ed elaborazione numerica dei segnali, eventualmente nella forma modificata (5.2.2)), ma si rivela poco conveniente quando la sequenza $x[n]$ proviene da un'operazione di campionamento di un segnale analogico $x(t)$. La presenza della frequenza *normalizzata* infatti non consente di stabilire un legame immediato con la frequenza (misurata in Hz) delle componenti nella trasformata $X(f)$ del segnale analogico di partenza. Se allora il periodo di campionamento è T, si può definire la variabile

$$f \triangleq \frac{F}{T} = F \cdot f_c \tag{5.2.4}$$

ottenuta "denormalizzando" la frequenza normalizzata F, e avente le dimensioni di una *frequenza*. Sostituendo nella definizione originaria di trasformata (5.2.1) l'espressione della frequenza normalizzata $F = f\,T$ ottenuta dalla (5.2.4), otteniamo una *nuova* definizione di trasformata di Fourier di una sequenza, che risulta adesso funzione complessa della frequenza f *misurabile in Hz*:

$$\overline{X}(f) \stackrel{\Delta}{=} \sum_{n=-\infty}^{\infty} x[n]\, e^{-j2\pi nfT} = \mathcal{F}\left[x[n]\right] \tag{5.2.5}$$

In questa definizione, che adotteremo definitivamente, abbiamo introdotto un *soprassegno* al nome della trasformata per distinguere facilmente la trasformata di una sequenza $\overline{X}(f)$ dalla trasformata di un segnale continuo $X(f)$.

Poiché la funzione $\overline{X}(F)$ è periodica di periodo 1, la funzione $\overline{X}(f)$ è periodica in ambito frequenziale di un periodo pari alla *frequenza di campionamento* $f_c = 1/T$:

$$\overline{X}(f) = \overline{X}\left(f + \frac{1}{T}\right) \tag{5.2.6}$$

Questa periodicità è caratteristica delle trasformate delle sequenze, e diventa chiara se introduciamo la relazione, inversa della (5.2.5), di *antitrasformazione*. Quest'ultima permette di esprimere la sequenza $x[n]$ attraverso un integrale di Fourier (o *antitrasformata di Fourier di una sequenza*) analogo alla (3.1.8):

$$x[n] = T \int_{-1/2T}^{1/2T} \overline{X}(f)\, e^{j2\pi nfT}\, df \tag{5.2.7}$$

Per giustificare questa relazione, riprendiamo in considerazione la definizione di trasformata (5.2.5), moltiplichiamo ambo i membri per un'oscillazione complessa alla frequenza f, e integriamo sull'intervallo frequenziale $[-1/2T, 1/2T]$:

$$\int_{-1/2T}^{1/2T} \overline{X}(f)\, e^{j2\pi nfT} df = \int_{-1/2T}^{1/2T} \sum_{m=-\infty}^{\infty} x[m]\, e^{-j2\pi mfT} e^{j2\pi nfT} df$$

$$= \sum_{m=-\infty}^{\infty} x[m] \int_{-1/2T}^{1/2T} e^{-j2\pi(m-n)fT} df \qquad (5.2.8)$$

Se si osserva che

$$\int_{-1/2T}^{1/2T} e^{-j2\pi(m-n)fT} df = \begin{cases} 0 & m \neq n \\ 1/T & m = n \end{cases} \qquad (5.2.9)$$

la serie a secondo membro della (5.2.8) si riduce all'unico termine per cui $m = n$, e cioè la (5.2.8) diventa la (5.2.7).

Il significato dell'espansione (5.2.7) della sequenza $x[n]$ è quello consueto per l'analisi di Fourier: il segnale dato viene espresso come *sovrapposizione* di un continuo di componenti frequenziali di ampiezza e fase regolate dall'andamento di $\overline{X}(f)$. Mentre un segnale analogico necessita di componenti a *tutte* le frequenze sull'asse reale da $-\infty$ a $+\infty$, cioè in un ambito *illimitato*, per esprimere una sequenza in ambito frequenziale sono sufficienti le sole componenti aventi frequenze comprese nell'intervallo *limitato* $[-1/2T, 1/2T]$. In un certo modo, questo risultato è pienamente giustificato dalla *periodicità* della trasformata di una sequenza, nel senso che le sole componenti veramente significative sono quelle nel periodo base. Anche per la trasformata di una sequenza è comunque d'uso introdurre lo spettro di ampiezza $\overline{A}(f) = |\overline{X}(f)|$ e lo spettro di fase $\overline{\theta}(f) = \angle\overline{X}(f)$.

La periodicità della trasformata di una sequenza, e/o la limitatezza dell'intervallo frequenziale significativo nello spettro del segnale, sono conseguenze di un unico fenomeno: nell'ambito dei segnali a tempo discreto, due oscillazioni sinusoidali (complesse) rispettivamente alla frequenza f_0 e $f_0 + m/T$ (m intero) sono *indistinguibili*. Infatti:

$$\exp[j2\pi(f_0 + m/T)nT] = \exp(j2\pi f_0 nT)\exp(j2\pi mn) = \exp(j2\pi f_0 nT) \qquad (5.2.10)$$

La Figura 5.8 rappresenta un esempio di questa situazione per il caso particolare $f_0 = 0$ e $m = 1$. Quest'esempio conferma che non ha senso pensare a componenti significative nello spettro del segnale discreto al di fuori di un intervallo-

base di ampiezza pari alla frequenza di campionamento, poiché tali componenti sono *repliche indistinguibili* di quelle già presenti nel periodo-base. Ma questa è proprio, in altre parole, la definizione di *periodicità* della trasformata $\overline{X}(f)$!

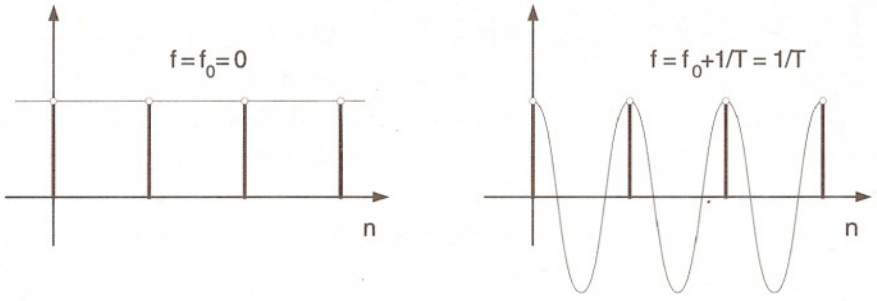


Figura 5.8 Oscillazioni a tempo discreto a frequenza 0 e $1/T$

Per la serie (5.2.5) si possono porre dei problemi di convergenza. Una condizione sufficiente per l'esistenza della trasformata è *l'assoluta sommabilità* della sequenza, cioè la condizione

$$\sum_{n=-\infty}^{\infty} |x[n]| < +\infty \tag{5.2.11}$$

Infatti, essendo

$$|\overline{X}(f)| = \left| \sum_{n=-\infty}^{\infty} x[n]\, e^{-j2\pi nfT} \right| \leq \sum_{n=-\infty}^{\infty} |x[n]| \tag{5.2.12}$$

il verificarsi della (5.2.11) assicura la convergenza della serie per tutti i valori della frequenza.

Esempio 5.1

Calcoliamo la trasformata di Fourier della sequenza $\delta[n]$. Essa è data da

$$\delta[n] \Leftrightarrow \overline{\Delta}(f) = \sum_{n=-\infty}^{\infty} \delta[n]\, e^{-j2\pi nfT} = 1 \tag{E5.1.1}$$

poiché nella serie l'unico termine non nullo è quello per $n = 0$. La trasformata $\overline{\Delta}(f)$ è quindi una funzione *periodica* di periodo $1/T$ che, in ciascun periodo, assume un valore costante e pari a 1 (Figura 5.9).

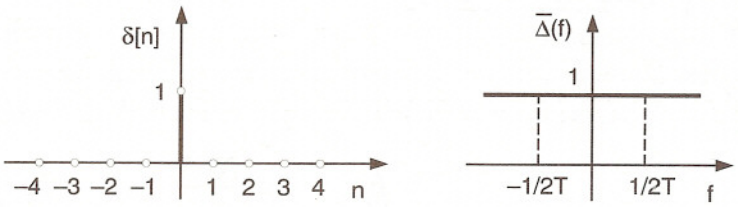


Figura 5.9 Trasformata della sequenza $\delta[n]$

❑

Esempio 5.2

Calcoliamo la trasformata di Fourier dell'*impulso rettangolare discreto* di Figura 5.7

$$x[n]=u[n]-u[n-N] \tag{E5.2.1}$$

Applicando la definizione,

$$\bar{X}(f)=\sum_{n=-\infty}^{+\infty}x[n]\,e^{-j2\pi nfT}=\sum_{n=0}^{N-1}e^{-j2\pi nfT}=\sum_{n=0}^{N-1}\left[e^{-j2\pi fT}\right]^n \tag{E5.2.2}$$

e ricordando che

$$\sum_{n=0}^{N-1}q^n=\frac{1-q^N}{1-q} \tag{E5.2.3}$$

si trova

$$\bar{X}(f)=\frac{1-e^{-j2\pi NfT}}{1-e^{-j2\pi fT}}=\frac{e^{-j\pi NfT}}{e^{-j\pi fT}}\frac{e^{j\pi NfT}-e^{-j\pi NfT}}{e^{j\pi fT}-e^{-j\pi fT}}=e^{-j\pi(N-1)fT}\frac{\sin(N\pi fT)}{\sin(\pi fT)} \tag{E5.2.4}$$

Lo spettro di ampiezza è allora

$$\left|\bar{X}(f)\right|=\left|\frac{\sin(N\pi fT)}{\sin(\pi fT)}\right| \tag{E5.2.5}$$

ed è rappresentato in Figura 5.10a per $N=10$ nell'intervallo $f\in[-1/T,1/T]$ per evidenziarne la *periodicità*. L'andamento dello spettro di ampiezza in vicinanza di $f=0$ ricorda quello di una funzione $N\,\mathrm{sinc}(NfT)$. Inoltre, restringendoci all'intervallo "base" $f\in[-1/2T,1/2T]$, esso si annulla per le frequenze

$$f_k = \frac{k}{NT} \quad , \quad k = \pm 1, \pm 2, \ldots, \pm N/2 \tag{E5.2.6}$$

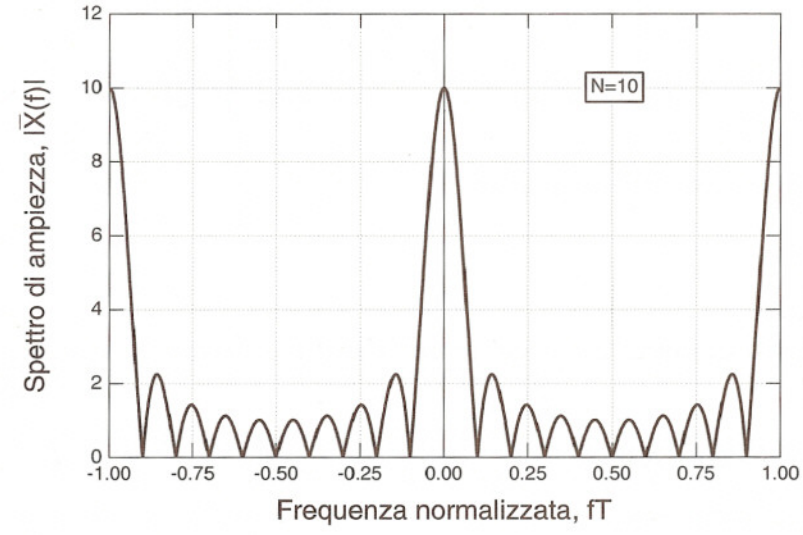


Figura 5.10 Spettro di ampiezza (a) e fase (b) dell'impulso rettangolare discreto

e assume per $f = 0$ il valore massimo dato da

$$\left|\overline{X}(0)\right| = \lim_{f \to 0} \left|\overline{X}(f)\right| = N \tag{E5.2.7}$$

L'andamento dello spettro di fase $\angle\overline{X}(f)$ nell'intervallo $f \in [-1/T, 1/T]$ è quello di Figura 5.10b, disegnato ancora per $N = 10$. L'andamento lineare è dovuto al posizionamento non simmetrico dell'impulso rispetto a n=0 (si veda il teorema del ritardo 5.3.2), cioè al termine $e^{-j\pi(N-1)fT}$ nella trasformata. ❑

Esempio 5.3

Calcoliamo la trasformata di Fourier della sequenza *esponenziale discreto*:

$$x[n] = a^n u[n] \tag{E5.3.1}$$

Si ha che

$$\overline{X}(f) = \sum_{n=-\infty}^{+\infty} x[n]\, e^{-j2\pi nfT} = \sum_{n=0}^{+\infty} a^n e^{-j2\pi nfT} = \sum_{n=0}^{+\infty} \left[a\, e^{-j2\pi fT}\right]^n \tag{E5.3.2}$$

Ricordando la formula della somma della *serie geometrica*,

$$\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q} \quad , \quad |q| < 1 \tag{E5.3.3}$$

si trova immediatamente

$$\overline{X}(f) = \frac{1}{1 - a\, e^{-j2\pi fT}} \quad , \quad |a| < 1 \tag{E5.3.4}$$

Lo spettro di ampiezza della sequenza data è allora (Figura 5.11a)

$$|\overline{X}(f)| = \frac{1}{\sqrt{1 + a^2 - 2\, a \cos(2\pi fT)}} \tag{E5.3.5}$$

mentre il suo spettro di fase è dato da (Figura 5.11b)

$$\angle\overline{X}(f) = \text{arctg}\left[\frac{a \sin(2\pi fT)}{1 - a \cos(2\pi fT)}\right] \tag{E5.3.6}$$

❑

Le proprietà della trasformata di Fourier di una sequenza (simmetria Hermitiana ecc.) relativamente a quelle del segnale temporale (segnale reale, segnale pari o dispari ecc.) sono sostanzialmente identiche a quelle ricavate nel Paragrafo 3.2 a proposito della trasformata dei segnali analogici, e non verranno

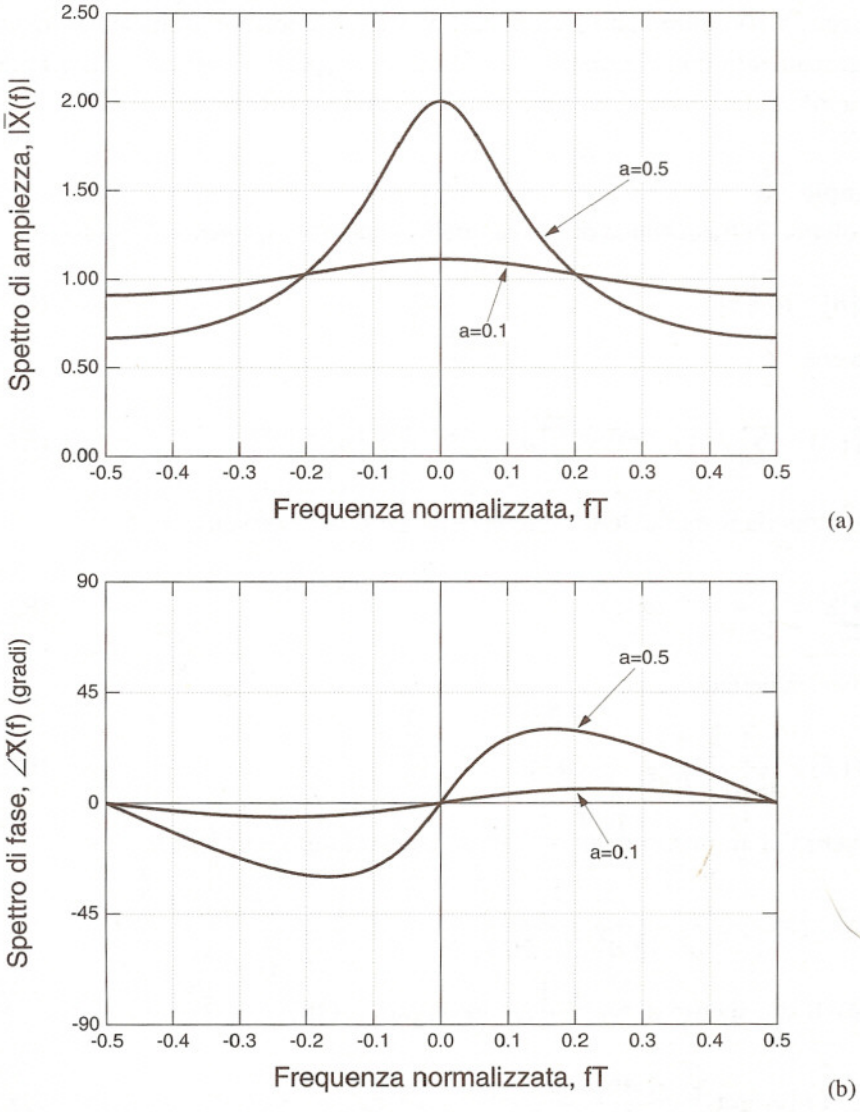


Figura 5.11 Spettro di ampiezza (a) e fase (b) della sequenza esponenziale discreto

ulteriormente ribadite. Sulla base delle relazioni di *analisi* (5.2.5) e di *sintesi* (5.2.7), che di seguito riportiamo, il lettore è in grado di ricavare tali proprietà autonomamente. Alcune differenze rispetto al caso del tempo continuo si riscontrano invece nelle proprietà analoghe a quelle viste nel Paragrafo 3.3.

$$\boxed{\begin{aligned} x[n] &= T \int_{-1/2T}^{1/2T} \overline{X}(f)\, e^{j2\pi nfT} df \\ \overline{X}(f) &= \sum_{n=-\infty}^{\infty} x[n] e^{-j2\pi nfT} \end{aligned}} \quad \text{ovvero} \quad x[n] \Leftrightarrow \overline{X}(f) \tag{5.2.13}$$

5.3 Teoremi sulla trasformata di Fourier di una sequenza

Elenchiamo e dimostriamo invece alcune proprietà fondamentali della trasformata di Fourier di una sequenza, spesso analoghe, ma in qualche caso significativamente differenti da quelle per i segnali a tempo continuo esaminate nel Paragrafo 3.3. Tali proprietà saranno indicate come *teoremi*.

5.3.1 Teorema di *linearità*

Supponiamo che una sequenza $x[n]$ sia espressa come combinazione lineare delle sequenze $x_1[n]$ e $x_2[n]$:

$$x[n] = a\, x_1[n] + b\, x_2[n] \tag{5.3.1}$$

con a e b costanti. Se si indicano con $\overline{X}_1(f)$ e $\overline{X}_2(f)$ le trasformate rispettivamente delle sequenze $x_1[n]$ e $x_2[n]$, la trasformata $\overline{X}(f)$ di $x[n]$ è

$$\overline{X}(f) = a\, \overline{X}_1(f) + b\, \overline{X}_2(f) \tag{5.3.2}$$

Questa proprietà è una banale conseguenza della definizione di trasformata di Fourier di una sequenza, e si dimostra in modo analogo a quanto visto nel Paragrafo 3.3.

5.3.2 Teorema del *ritardo*

Consideriamo una sequenza $x[n]$ con trasformata $\overline{X}(f)$. La trasformata della sequenza $x[n-k]$ ottenuta ritardando $x[n]$ di k passi è espressa da

$$x[n-k] \Leftrightarrow \overline{X}(f)\, e^{-j2\pi kfT} \tag{5.3.3}$$

Per dimostrare questa proprietà basta osservare che

$$\begin{aligned} \mathcal{F}[x[n-k]] &= \sum_{n=-\infty}^{\infty} x[n-k]\, e^{-j2\pi nfT} = \sum_{m=-\infty}^{\infty} x[m]\, e^{-j2\pi(m+k)fT} \\ &= e^{-j2\pi kfT} \sum_{m=-\infty}^{\infty} x[m]\, e^{-j2\pi mfT} = \overline{X}(f)\, e^{-j2\pi kfT} \end{aligned} \tag{5.3.4}$$

avendo effettuato negli sviluppi il cambiamento di variabile $m = n - k$.

5.3.3 Teorema della *modulazione*

La trasformata della sequenza $x[n]\,e^{j2\pi nf_0T}$, ottenuta "modulando" $x[n]$ con la sequenza $e^{j2\pi nf_0T}$, è espressa da

$$x[n]\,e^{j2\pi nf_0T} \Leftrightarrow \overline{X}(f-f_0) \tag{5.3.5}$$

Infatti si ha:

$$\mathcal{F}\left[x[n]\,e^{j2\pi nf_0T}\right] = \sum_{n=-\infty}^{\infty} x[n]\,e^{j2\pi nf_0T}e^{-j2\pi nfT}$$

$$= \sum_{n=-\infty}^{\infty} x[n]\,e^{-j2\pi n(f-f_0)T} = \overline{X}(f-f_0) \tag{5.3.6}$$

Poiché la funzione $\overline{X}(f)$ è periodica di periodo $1/T$, la funzione $\overline{X}(f-f_0)$, ottenuta traslando $\overline{X}(f)$ della quantità f_0, coincide con la funzione $\overline{X}(f-|f_0|_{1/T})$, ottenuta invece traslando $\overline{X}(f)$ della quantità

$$|f_0|_{1/T} \triangleq f_0 - m/T \quad , \quad m = \mathrm{int}\left(f_0/(1/T)\right) \tag{5.3.7}$$

cioè f_0 *modulo* $1/T$. Questa è un'ulteriore conferma della proprietà (5.2.10) delle oscillazioni complesse a tempo discreto.

5.3.4 Teorema della *somma di convoluzione*

Prima di introdurre questa proprietà, definiamo la sequenza $z[n]$ *somma di convoluzione* tra le sequenze aperiodiche $x[n]$ e $y[n]$:

$$z[n] = x[n] \otimes y[n] \triangleq \sum_{k=-\infty}^{+\infty} x[k]\,y[n-k] = \sum_{k=-\infty}^{+\infty} y[k]\,x[n-k] \tag{5.3.8}$$

La somma di convoluzione gode naturalmente delle stesse proprietà *commutativa*, *associativa* e *distributiva* già dimostrate nel Paragrafo 3.3 per l'integrale di convoluzione tra segnali analogici.

Ora, il teorema della somma di convoluzione afferma che la trasformata di Fourier della sequenza $z[n]$ è data dal prodotto delle trasformate delle sequenze $x[n]$ e $y[n]$:

$$z[n] = x[n] \otimes y[n] \Leftrightarrow \overline{Y}(f)\,\overline{X}(f) = \overline{Z}(f) \tag{5.3.9}$$

Infatti

$$\overline{Z}(f) = \sum_{n=-\infty}^{\infty}\sum_{k=-\infty}^{\infty} x[k]\,y[n-k]\,e^{-j2\pi nfT} = \sum_{k=-\infty}^{\infty} x[k]\sum_{n=-\infty}^{\infty} y[n-k]\,e^{-j2\pi nfT} \tag{5.3.10}$$

avendo invertito l'ordine delle due sommatorie. Se si osserva che (per il teorema del ritardo)

$$\sum_{n=-\infty}^{\infty} y[n-k]\, e^{-j2\pi nfT} = \overline{Y}(f)\, e^{-j2\pi kfT} \tag{5.3.11}$$

la relazione (5.3.10) può essere riscritta nella forma

$$\overline{Z}(f) = \sum_{k=-\infty}^{\infty} x[k]\, \overline{Y}(f)\, e^{-j2\pi kfT} = \overline{Y}(f) \sum_{k=-\infty}^{\infty} x[k]\, e^{-j2\pi kfT} = \overline{Y}(f)\, \overline{X}(f) \tag{5.3.12}$$

L'importanza di questo risultato diventerà chiara nello studio dei sistemi lineari stazionari a tempo discreto (Capitolo 6).

5.3.5 Teorema del *prodotto*

Consideriamo adesso la sequenza $p[n]$ data dal prodotto fra la sequenza $x[n]$ e la sequenza $y[n]$

$$p[n] = x[n] \cdot y[n] \tag{5.3.13}$$

e calcoliamone la trasformata di Fourier:

$$\begin{aligned} \overline{P}(f) &= \sum_{n=-\infty}^{\infty} p[n]\, e^{-j2\pi nfT} = \sum_{n=-\infty}^{\infty} x[n]\, y[n]\, e^{-j2\pi nfT} \\ &= \sum_{n=-\infty}^{\infty} T \int_{-1/2T}^{1/2T} \overline{X}(\nu)\, e^{j2\pi n\nu T} d\nu\; y[n]\, e^{-j2\pi nfT} \end{aligned} \tag{5.3.14}$$

avendo espresso $x[n]$ come antitrasformata di Fourier di $\overline{X}(f)$. In questo passaggio, è stata usata una variabile "muta" ν nell'operazione di antitrasformazione per non creare ambiguità con la variabile f da cui dipende la trasformata $\overline{P}$.

Se nella (5.3.14) si inverte l'ordine delle operazioni di somma e di integrazione si ricava

$$\overline{P}(f) = T \int_{-1/2T}^{1/2T} \overline{X}(\nu) \sum_{n=-\infty}^{\infty} y[n]\, e^{-j2\pi n(f-\nu)T} d\nu = T \int_{-1/2T}^{1/2T} \overline{X}(\nu)\overline{Y}(f-\nu) d\nu \tag{5.3.15}$$

che permette di stabilire la relazione

$$p[n] = x[n]\, y[n] \Leftrightarrow \overline{P}(f) = T \int_{-1/2T}^{1/2T} \overline{X}(\nu)\overline{Y}(f-\nu) d\nu \tag{5.3.16}$$

Poiché le funzioni $\overline{X}(v)$ e $\overline{Y}(f-v)$ sono periodiche di periodo $1/T$, l'integrazione nella (5.3.16) può essere svolta su di un qualsiasi intervallo frequenziale di ampiezza $1/T$.

L'integrale a secondo membro della (5.3.16) rappresenta la cosiddetta *convoluzione ciclica* o *periodica* fra la trasformate $\overline{X}(f)$ e $\overline{Y}(f)$. La convoluzione ciclica è un'operazione che si definisce tra funzioni *periodiche* come le trasformate delle sequenze. Si nota che la funzione integranda è analoga a quella che si ha nella convoluzione cosiddetta *lineare* o *aperiodica* (3.3.43) eseguita tra funzioni *aperiodiche*, ma l'integrale viene calcolato su di un solo periodo, e il risultato viene diviso per l'ampiezza del periodo stesso, $1/T$.

5.3.6 Teorema dell'*incremento*

La derivata di un segnale a tempo continuo $x(t)$ all'istante $t=nT$ può essere approssimata con il seguente rapporto incrementale:

$$\left.\frac{dx(t)}{dt}\right|_{t=nT} \cong \frac{x(nT)-x(nT-T)}{T} = \frac{x[n]-x[n-1]}{T} \tag{5.3.17}$$

avendo definito la sequenza $x[n] \triangleq x(nT)$ ottenuta per campionamento dal segnale continuo $x(t)$. Se si introduce allora l'operatore *incremento* Δ definito dalla relazione

$$\Delta x[n] \triangleq x[n]-x[n-1] \tag{5.3.18}$$

è ragionevole immaginare la sequenza $\Delta x[n]$ degli incrementi di $x[n]$ come una sorta di omologo a tempo discreto della derivata del segnale a tempo continuo $dx(t)/dt$. Utilizzando il teorema del ritardo si può scrivere:

$$\Delta x[n] \Leftrightarrow \overline{X}(f)-\overline{X}(f)e^{-j2\pi fT} = \overline{X}(f)\left(1-e^{-j2\pi fT}\right) \tag{5.3.19}$$

5.3.7 Teorema della *sequenza somma*

Consideriamo adesso la *sequenza somma* $y[n]$ di una sequenza data $x[n]$:

$$y[n] \triangleq \sum_{k=-\infty}^{n} x[k] \tag{5.3.20}$$

Vogliamo dimostrare che la trasformata della sequenza $y[n]$ è espressa da

$$\overline{Y}(f) = \frac{\overline{X}(f)}{1-e^{-j2\pi fT}} \tag{5.3.21}$$

purché $\overline{X}(0) = 0$. Il teorema dell'incremento permette infatti di scrivere:

$$\Delta y[n] \Leftrightarrow \overline{Y}(f)\left(1 - e^{-j2\pi fT}\right) \tag{5.3.22}$$

D'altronde

$$\Delta y[n] = y[n] - y[n-1] = \sum_{k=-\infty}^{n} x[k] - \sum_{k=-\infty}^{n-1} x[k] = x[n] \tag{5.3.23}$$

e quindi

$$\overline{X}(f) = \overline{Y}(f)\left(1 - e^{-j2\pi fT}\right) \tag{5.3.24}$$

Ciò permette di concludere che:

$$y[n] = \sum_{k=-\infty}^{n} x[k] \Leftrightarrow \overline{Y}(f) = \frac{\overline{X}(f)}{1 - e^{-j2\pi fT}} \tag{5.3.25}$$

Se però si considera la (5.3.24) per $f = 0$, si ha

$$\overline{X}(0) = \overline{Y}(0)\left(1 - e^{-j0}\right) \tag{5.3.26}$$

che non può essere valida se $\overline{X}(0) \neq 0$. Quindi, condizione per l'applicabilità del teorema nella forma (5.3.25) è che valga

$$\overline{X}(0) = \sum_{n=-\infty}^{\infty} x[n] = 0 \tag{5.3.27}$$

5.4 La condizione di Nyquist e il teorema del campionamento

5.4.1 La condizione di Nyquist

Riprendiamo in considerazione il campionamento di un segnale a tempo continuo $x(t)$:

$$x[n] = x(nT) \tag{5.4.1}$$

e cerchiamo di determinare le conseguenze *in ambito frequenziale* di questa relazione valida in ambito temporale. Indichiamo come di consueto con $\overline{X}(f)$ e $X(f)$ rispettivamente la trasformata di Fourier della sequenza $x[n]$ e del segnale analogico $x(t)$.

Ovviamente si ha:

$$\overline{X}(f)=\sum_{n=-\infty}^{+\infty}x[n]\,e^{-j2\pi nfT}=\sum_{n=-\infty}^{+\infty}x(nT)\,e^{-j2\pi nfT} \tag{5.4.2}$$

Esprimiamo adesso i *campioni* del segnale a tempo continuo $x(t)$ attraverso l'integrale di Fourier (3.1.8), valutato naturalmente all'istante $t=nT$:

$$\overline{X}(f)=\sum_{n=-\infty}^{\infty}\left(\int_{-\infty}^{\infty}X(v)\,e^{j2\pi vnT}dv\right)e^{-j2\pi nfT}=\int_{-\infty}^{\infty}X(v)\sum_{n=-\infty}^{\infty}e^{-j2\pi n(f-v)T}dv \tag{5.4.3}$$

avendo effettuato lo scambio dell'ordine di somma e di integrazione. Per semplificare la relazione (5.4.3) riconsideriamo lo sviluppo in serie di Fourier (E3.17.4) del segnale "pettine di δ":

$$\sum_{n=-\infty}^{+\infty}\delta(t-nT)=\frac{1}{T}\sum_{k=-\infty}^{\infty}e^{j\frac{2\pi kt}{T}} \tag{5.4.4}$$

Calcolando la trasformata di Fourier dei due membri di questa relazione si ha

$$\sum_{n=-\infty}^{\infty}e^{-j2\pi nfT}=\frac{1}{T}\sum_{k=-\infty}^{+\infty}\delta\left(f-\frac{k}{T}\right) \tag{5.4.5}$$

e sostituendo questo risultato nella (5.4.3) si ha

$$\begin{aligned}\overline{X}(f)&=\int_{-\infty}^{+\infty}X(v)\frac{1}{T}\sum_{k=-\infty}^{+\infty}\delta\left(f-v-\frac{k}{T}\right)dv\\&=\frac{1}{T}\sum_{k=-\infty}^{+\infty}\int_{-\infty}^{+\infty}X(v)\,\delta\left(v-\left(f-\frac{k}{T}\right)\right)dv\end{aligned} \tag{5.4.6}$$

Sfruttando infine la proprietà campionatrice della funzione δ si ottiene

$$\overline{X}(f)=\frac{1}{T}\sum_{k=-\infty}^{\infty}X\left(f-\frac{k}{T}\right) \tag{5.4.7}$$

che rappresenta la relazione cercata (si veda anche la (3.5.9)).

Questa relazione mostra che la trasformata di Fourier di una sequenza ottenuta per campionamento si ricava come *periodicizzazione* della trasformata del segnale analogico di partenza, con un periodo di ripetizione in frequenza pari alla *frequenza di campionamento* $f_c=1/T$. Un esempio significativo è illustrato nelle Figure 5.12a-c: lo spettro del segnale $x(t)$ è rappresentato in Figura 5.12a mentre lo spettro della sequenza $x[n]$ è rappresentato nelle Figure 5.12b-c per due scelte diverse della frequenza di campionamento.

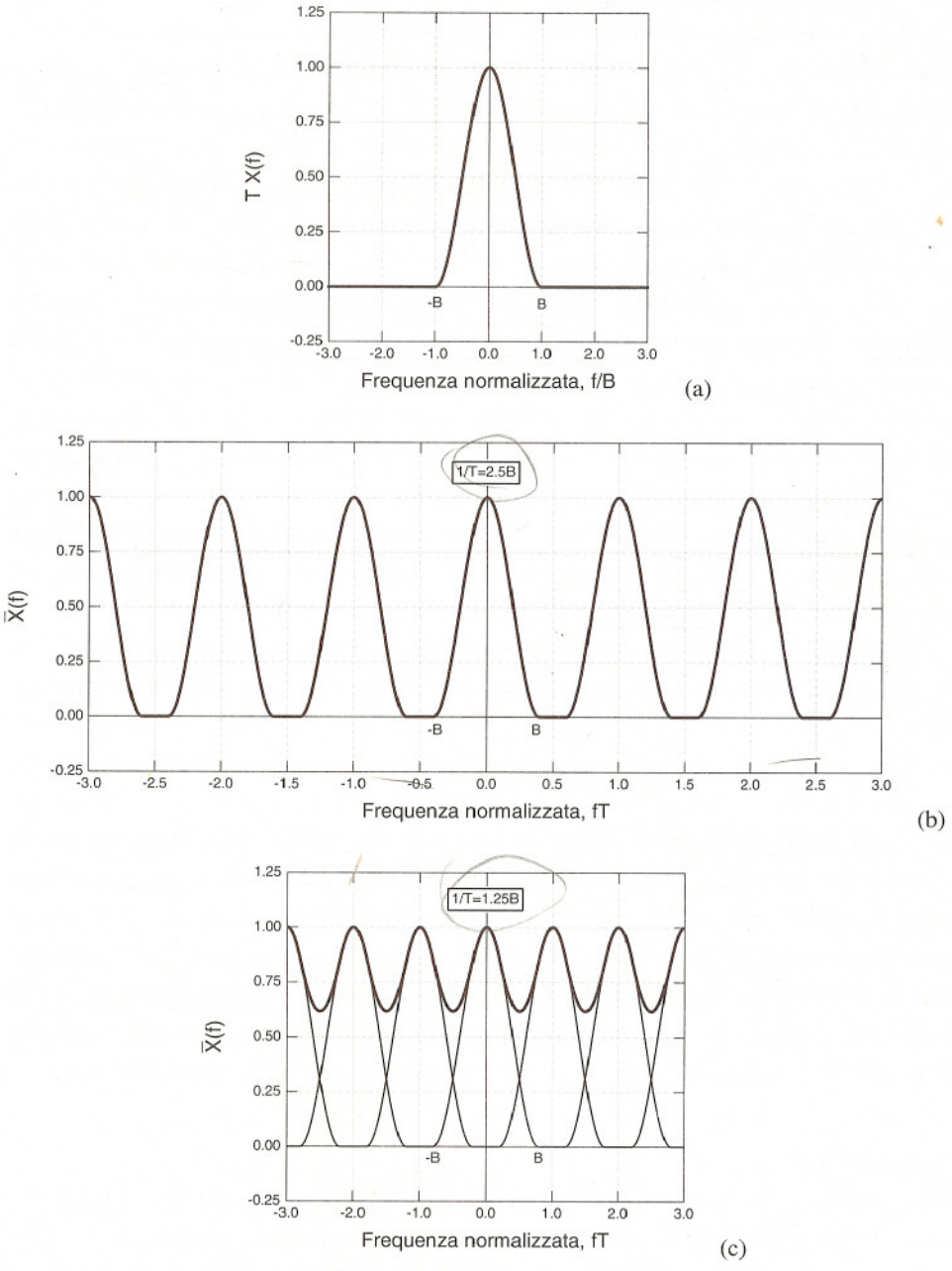


Figura 5.12 Trasformata del segnale analogico $x(t)$ (a) e della sequenza $x[n]$ con frequenza di campionamento $5B/2$ (b) e $5B/4$ (c)

In particolare, il segnale di partenza ha *banda B rigorosamente limitata*, e le due

frequenze di campionamento sono pari rispettivamente a $2.5B$ nel caso (b), e $1.25B$ nel caso (c). Come si nota, c'è una differenza sostanziale nelle trasformate della sequenza $x[n]$ nei due casi delle Figure 5.12b e 5.12c. In quest'ultima, la frequenza di campionamento è tale che le varie *repliche* della trasformata di $x(t)$ centrate sui multipli della frequenza di campionamento e derivanti dalla periodicizzazione dello spettro vengono a *sovrapporsi*. Nel caso di Figura 5.12b, invece, la frequenza di campionamento è sufficientemente alta, e non si ha sovrapposizione. In quest'ultima situazione, il periodo frequenziale base $[-1/2T, 1/2T]$ contiene una *replica* non distorta della trasformata $X(f)$ del segnale a tempo continuo originario. Viceversa, nell'altro caso le varie repliche dello spettro "interferiscono" sommandosi alla replica base; nell'ambito dell'intervallo $[-1/2T, 1/2T]$ quest'ultima è accompagnata dal cosiddetto *errore di aliasing* creato dalle repliche (alias) dello spettro-base, che porta a una *distorsione* del segnale.

Se il segnale è a banda limitata, dunque, è possibile trovare una condizione che garantisce *assenza di aliasing*: la banda del segnale analogico di partenza B deve essere più piccola dell'estremo superiore dell'intervallo "base" $[-1/2T, 1/2T]$, cioè $B \leq 1/2T$. In altri termini, fissata la banda del segnale B, la frequenza di campionamento deve essere scelta in modo che valga la condizione

$$f_c = \frac{1}{T} \geq 2B \tag{5.4.8}$$

detta *condizione di Nyquist*. Nella Figura 5.12b la condizione di Nyquist è soddisfatta ($f_c = 2.5B > 2B$), e l'intervallo frequenziale base contiene una *replica indistorta* dello spettro del segnale analogico di partenza, cosa che non accade nella Figura 5.12c, in cui $f_c = 1.25B < 2B$. Questa osservazione sembra suggerire la possibilità, in assenza di aliasing, di *ricostruire il segnale originario (a banda limitata) elaborando il segnale campionato*. Nel paragrafo successivo illustreremo come questa operazione possa essere effettuata in teoria e in pratica.

Esempio 5.4

Osservazioni sperimentali di carattere fisiologico mostrano che l'orecchio umano può udire segnali costituiti da componenti frequenziali comprese al più nella banda tra 20 Hz e 20 kHz. Suoni con frequenza inferiore ai 20 Hz vengono percepiti come "vibrazioni" con tutto il corpo, mentre suoni con frequenza maggiore di 20 kHz sono inudibili e vengono chiamati *ultrasuoni*.

Possiamo quindi considerare un segnale audio come *rigorosamente limitato*

in banda con un limite di banda $B = 20$ kHz. Questa osservazione, insieme con la condizione di Nyquist, giustifica la scelta delle due frequenze standard di campionamento nell'elaborazione numerica dei segnali audio ad alta fedeltà: come già accennato nel Capitolo 1, i sistemi di registrazione su Compact-Disc adottano una $f_c = 44.1$ kHz, mentre i registratori DAT (Digital Audio Tape) hanno $f_c = 48$ kHz. Entrambi questi valori sono di poco superiori al limite minimo $f_c = 2B = 40$ kHz; viene introdotto un margine "di sicurezza" per facilitare (come vedremo in seguito) l'operazione di ricostruzione del segnale analogico.

Quando il segnale audio campionato deve essere *trasmesso*, come in un sistema di radiodiffusione, è importante cercare di ridurre al minimo il numero di campioni/s, cioè la frequenza di campionamento. Infatti, più campioni devono essere trasmessi nello stesso intervallo di tempo, maggiore deve essere la *capacità* (e quindi il costo) del sistema di trasmissione. Per questo motivo, in alcuni standard di radiodiffusione dell'audio digitale, si sceglie di campionare il segnale con $f_c = 32$ kHz. Questa frequenza *non* soddisfa chiaramente la condizione di Nyquist rispetto alla banda $B = 20$ kHz. Allora, per evitare problemi di aliasing, si antepone al convertitore A/D un *filtro anti-aliasing*, come mostrato in Figura 5.13, che limita convenientemente la banda del segnale analogico a un valore B' in modo da annullare l'aliasing per la frequenza di campionamento fissata. Nel caso dello standard con $f_c = 32$ kHz, il filtro anti-aliasing ha una banda $B' = 15$ kHz. È chiaro che in questo modo la qualità del segnale riprodotto sarà inferiore a quella dei sistemi CD e DAT per l'artificiale limitazione in banda, ma ancora sufficientemente elevata per una riproduzione godibile, e con una minore esigenza di capacità del sistema di trasmissione.

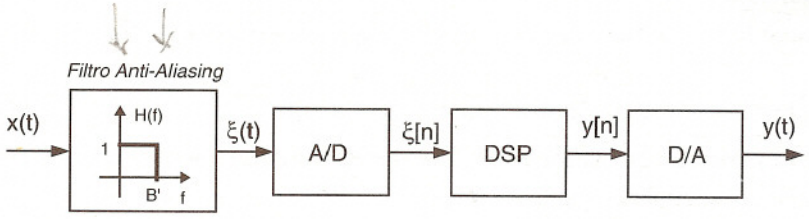


Figura 5.13 Elaborazione numerica del segnale con filtro anti-aliasing

❑

La condizione di Nyquist (5.4.8) pone dei vincoli sulla scelta della frequenza di campionamento se si desidera ricostruire un segnale a tempo continuo utilizzandone i campioni; in particolare, il periodo di campionamento deve essere

scelto in funzione della *banda* del segnale analogico. Gli esempi illustrati in Figura 5.14 giustificano ulteriormente questa condizione. Nella Figura 5.14a viene rappresentato un segnale $x(t)$ che ha una *rapidità di variazione* (e quindi una *banda*) molto maggiore di quella del segnale $y(t)$ di Figura 5.14b. Si intuisce allora che, per seguire con sufficiente accuratezza l'andamento del segnale, e quindi poter poi essere in grado di ricostruire il medesimo a partire dai campioni prelevati, si deve adottare un periodo di campionamento *più piccolo* (frequenza di campionamento maggiore) per il segnale $x(t)$ che per $y(t)$.

Negli esempi di Figura 5.14, il periodo di campionamento scelto è adeguato per $y(t)$, ma è palesemente troppo grande per $x(t)$. In questo senso, la frequenza di campionamento deve essere commisurata con la *banda* del segnale, come la condizione di Nyquist suggerisce.

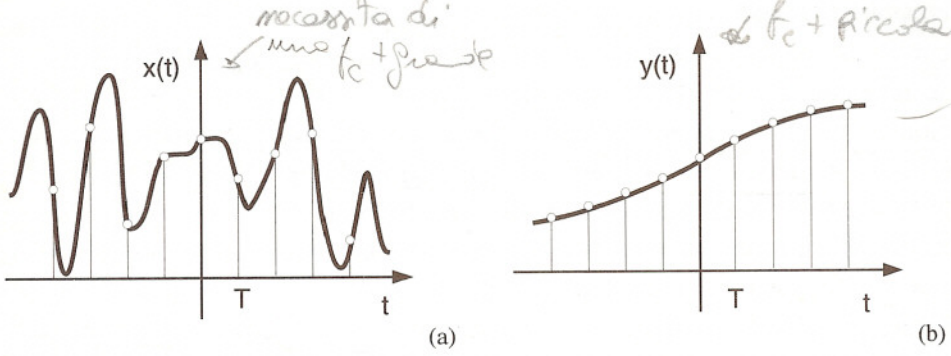


Figura 5.14 Esempi di campionamento di segnali a tempo continuo

Esempio 5.5

Ricaviamo la trasformata di Fourier della sequenza *costante*

$$x[n] = 1 \tag{E5.5.1}$$

Possiamo pensare $x[n]$ come risultante da un campionamento, con intervallo T arbitrario, del segnale costante a tempo continuo $x(t) = 1$. Poiché

$$x(t) = 1 \Leftrightarrow \delta(f) = X(f) \tag{E5.5.2}$$

applicando la relazione (5.4.7) del campionamento, si trova immediatamente

$$\overline{X}(f) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \delta\left(f - \frac{k}{T}\right) \tag{E5.5.3}$$

relazione rappresentata in Figura 5.15.

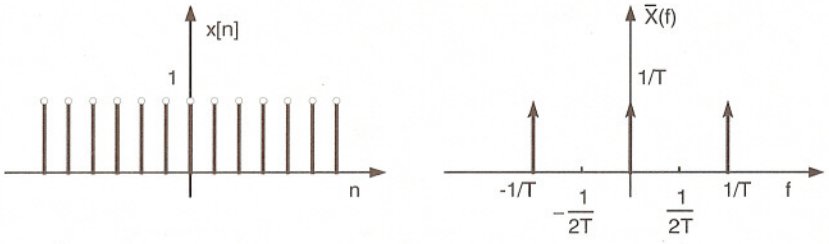


Figura 5.15 Trasformata di Fourier di una sequenza costante

❑

Esempio 5.6

Troviamo la trasformata di Fourier delle sequenze

$$x[n] = \cos(2\pi n f_0 T) \quad , \quad y[n] = \sin(2\pi n f_0 T) \tag{E5.6.1}$$

Dalla trasformata della sequenza costante dell'Esempio 5.5 e dal teorema della modulazione si ha che

$$\exp(j2\pi n f_0 T) \Leftrightarrow \frac{1}{T}\sum_{k=-\infty}^{\infty} \delta\left(f - f_0 - \frac{k}{T}\right) \tag{E5.6.2}$$

per cui, ricordando le formule di Eulero, si trova immediatamente

$$\begin{aligned} \overline{X}(f) &= \frac{1}{2T}\delta(f - |f_0|_{1/T}) + \frac{1}{2T}\delta(f + |f_0|_{1/T}) \\ \overline{Y}(f) &= \frac{1}{2jT}\delta(f - |f_0|_{1/T}) - \frac{1}{2jT}\delta(f + |f_0|_{1/T}) \end{aligned} \quad , \quad -\frac{1}{2T} \le f \le \frac{1}{2T} \tag{E5.6.3}$$

ove ci siamo limitati a considerare l'intervallo "base" della trasformata, senza esplicitamente indicare la periodicizzazione come nella (E5.6.2). ❑

Esempio 5.7

La trasformata di Fourier $\overline{X}(f)$ di una sequenza $x[n]$ è rappresentata in Figura 5.16. Determiniamo l'andamento della sequenza stessa. Attraverso la relazione di antitrasformazione si trova

$$x[n] = T\int_{-1/2T}^{1/2T} \overline{X}(f)\, e^{j2\pi n f T} df = T\int_{-B}^{B} e^{j2\pi n f T} df$$

$$= T\frac{e^{j2\pi nfT}}{j2\pi nT}\Bigg|_{-B}^{B} = 2BT\,\mathrm{sinc}(2BnT) \qquad \text{(E5.7.1)}$$

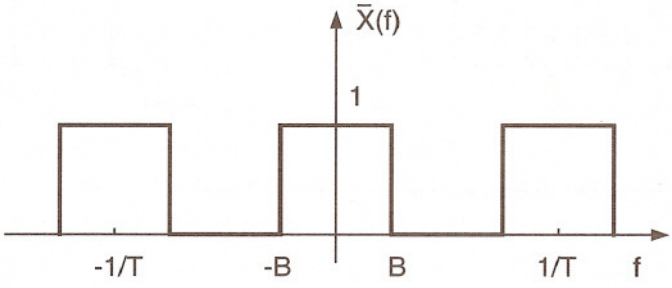


Figura 5.16 Trasformata di Fourier della sequenza $x[n]$ nell'Esempio 5.7

Allo stesso risultato si arriva più facilmente se si pensa la $\bar{X}(f)$ come derivante dalla *periodicizzazione* di una singola funzione rect(·):

$$\bar{X}(f) = \frac{1}{T}\sum_{k=-\infty}^{\infty} T \cdot \mathrm{rect}\left(\frac{f - k/T}{2B}\right) \qquad \text{(E5.7.2)}$$

Segue che la sequenza $x[n]$, rappresentata in Figura 5.17, può farsi derivare dal *campionamento* del segnale a tempo continuo

$$x(t) = \mathcal{F}^{-1}[T \cdot \mathrm{rect}(f/2B)] = 2BT\,\mathrm{sinc}(2Bt) \qquad \text{(E5.7.3)}$$

come suggerito del resto dalla (E5.7.1). ❑

5.4.2 Interpolazione a mantenimento

La ricostruzione di un segnale a tempo continuo a partire da una sequenza viene realizzata mediante un *interpolatore*. I vari tipi di interpolazione, che specificheremo con maggior dettaglio nel Paragrafo 5.4.3, possono in un certo senso considerarsi come una *generalizzazione* dell'operazione compiuta in pratica da un convertitore D/A per fornire in uscita un segnale a tempo continuo $\hat{x}(t)$ a partire dai valori (rappresentati su di un certo numero di cifre binarie) di una sequenza $x[n]$.

Lo schema di un sistema che esegue in pratica il campionamento del segnale e la successiva interpolazione, senz'alcuna elaborazione intermedia, è rappresentato in Figura 5.18 come cascata di due blocchi A/D e D/A, ovvero come successione di un campionatore ideale e di un *interpolatore a mantenimento.*

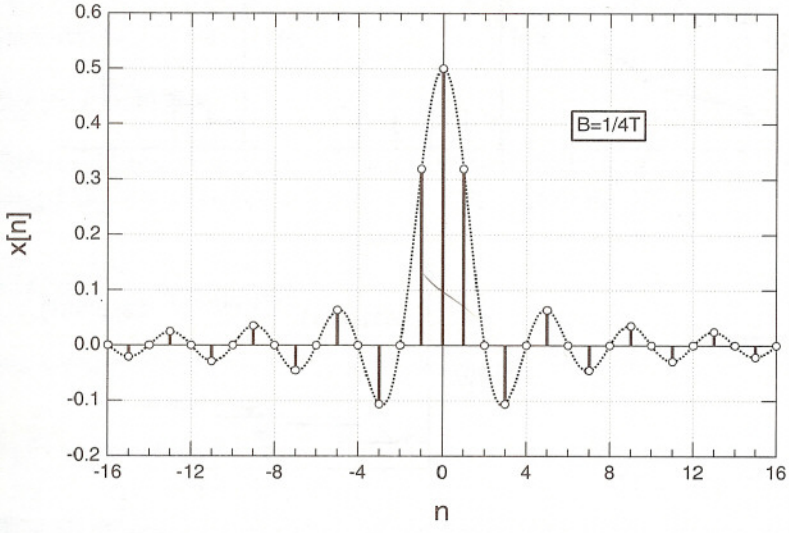


Figura 5.17 Sequenza $x[n]$ ottenuta per campionamento da $x(t)$ (Esempio 5.7)

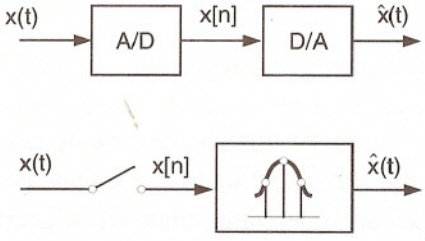


Figura 5.18 Campionamento e interpolazione a mantenimento

L'operazione svolta da quest'ultimo componente è in particolare raffigurata in Figura 5.19a: per costruire il segnale analogico di uscita, il valore n-esimo della sequenza d'ingresso $x[n]$ viene *mantenuto* a partire dall'istante nT e fino a che non sia disponibile (all'istante $(n+1)T$) il successivo valore $x[n+1]$.

Possiamo facilmente scrivere l'espressione del segnale interpolato $\hat{x}(t)$ in funzione dei valori della sequenza $x[n]$. La Figura 5.19b suggerisce che $\hat{x}(t)$ è costituito da una successione di *impulsi rettangolari* di durata T, applicati agli istanti nT e di ampiezza pari al relativo valore n-esimo della sequenza $x[n]$:

$$\hat{x}(t) = \sum_{n=-\infty}^{+\infty} x[n]\, p(t - nT) \tag{5.4.9}$$

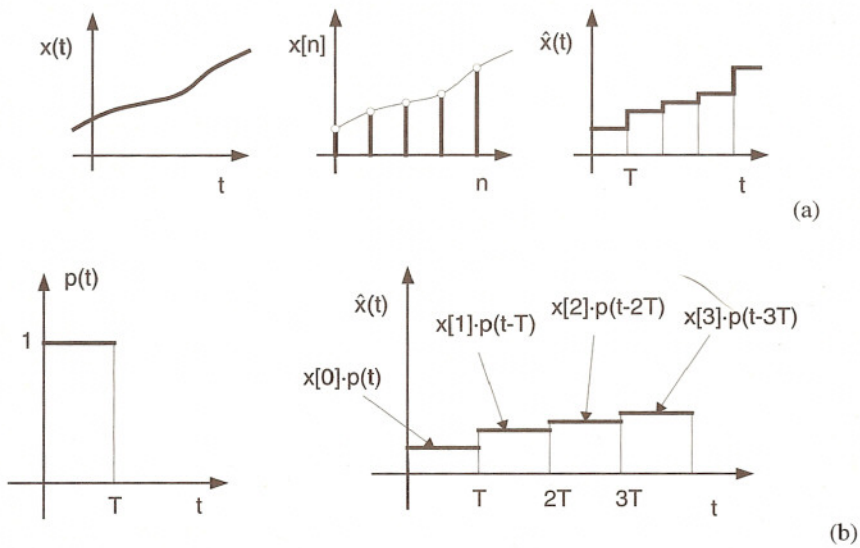


Figura 5.19 Campionamento e interpolazione a mantenimento

ove $p(t)$ è per l'appunto l'impulso rettangolare

$$p(t) = \text{rect}\left(\frac{t - T/2}{T}\right) \tag{5.4.10}$$

La Figura 5.19a mostra però chiaramente che il segnale ricostruito dall'interpolatore a mantenimento, che è una forma d'onda *costante a tratti,* non è una replica indistorta del segnale campionato $x(t)$. L'operazione che direttamente conduce da $x(t)$ al segnale costante a tratti $\hat{x}(t)$ viene indicata in elettronica con il nome di *Sample & Hold* (campiona e mantieni), abbreviato in S&H.

Cerchiamo dunque di comprendere più a fondo il comportamento dell'interpolatore a mantenimento, e di capire meglio le distorsioni che esso introduce rispetto al segnale di partenza, esaminandone il comportamento nel *dominio della frequenza*. Calcoliamo allora la trasformata di Fourier del segnale interpolato (5.4.9):

$$\hat{X}(f) = \sum_{n=-\infty}^{\infty} x[n]\, P(f)\, e^{-j2\pi nfT} = P(f) \sum_{n=-\infty}^{\infty} x[n]\, e^{-j2\pi nfT} = P(f)\, \overline{X}(f) \tag{5.4.11}$$

Questa relazione mostra che la trasformata di Fourier del segnale interpolato è data dal prodotto della trasformata continua dell'impulso di mantenimento $p(t)$

con la trasformata della sequenza $x[n]$. Essendo $p(t)$ espresso dalla (5.4.10), la sua trasformata $P(f)$ è

$$P(f) = T\,\mathrm{sinc}(fT)\,e^{-j\pi fT} \tag{5.4.12}$$

Abbiamo inoltre dimostrato che la trasformata della sequenza $x[n]$ è legata a quella del segnale a tempo continuo $x(t)$ dalla relazione

$$\overline{X}(f) = \frac{1}{T}\sum_{k=-\infty}^{\infty} X\left(f - \frac{k}{T}\right) \tag{5.4.13}$$

Allora, sostituendo le (5.4.12)-(5.4.13) nella (5.4.11) troviamo:

$$\hat{X}(f) = \mathrm{sinc}(fT)\,e^{-j\pi fT}\sum_{k=-\infty}^{\infty} X\left(f - \frac{k}{T}\right) \tag{5.4.14}$$

Per fare un esempio, supponiamo che la trasformata di Fourier del segnale $x(t)$ di banda B sia espressa da

$$X(f) = \frac{1}{2B}\cdot\frac{|f|}{B}\,\mathrm{rect}\left(\frac{f}{2B}\right) \tag{5.4.15}$$

mostrata in Figura 5.20, e che la frequenza di campionamento sia $1/T = 2.5B$, ovvero soddisfi la condizione di Nyquist. Utilizzando la relazione (5.4.14) appena ricavata si può rappresentare lo spettro di ampiezza $|\hat{X}(f)|$ di $\hat{x}(t)$. Tale spettro è illustrato nella Figura 5.21, in cui sono riportati anche, a linea rispettivamente grigia e tratteggiata, l'andamento dei due fattori che compongono $|\hat{X}(f)|$, e cioè $T\,|\overline{X}(f)|$ e $|\mathrm{sinc}(fT)|$.

Lo spettro del segnale ricostruito differisce apprezzabilmente da quello del segnale analogico di partenza in due aspetti fondamentali:

i) il segnale interpolato non è limitato in banda: l'operazione di ricostruzione del segnale introduce delle componenti frequenziali che non sono presenti nel segnale analogico $x(t)$. Esse derivano dalla presenza delle repliche dello spettro del segnale di partenza a cavallo dei multipli della frequenza di campionamento. Questi residui delle repliche sono chiamati *immagini*;

ii) anche all'interno della banda "utile", o meglio all'interno dell'intervallo base $[-1/2T, 1/2T]$, lo spettro del segnale ricostruito differisce dallo spettro del segnale di partenza. In assenza di aliasing, in tale intervallo i due spettri sono legati dalla relazione (che si deduce immediatamente dalla (5.4.14))

$$\hat{X}(f) = \frac{P(f)\,X(f)}{T} = X(f)\,\mathrm{sinc}(fT)\,e^{-j\pi fT} \quad , \quad -\frac{1}{2T} \le f \le \frac{1}{2T} \tag{5.4.16}$$

e quindi il segnale $x(t)$ subisce una *distorsione di ampiezza*.

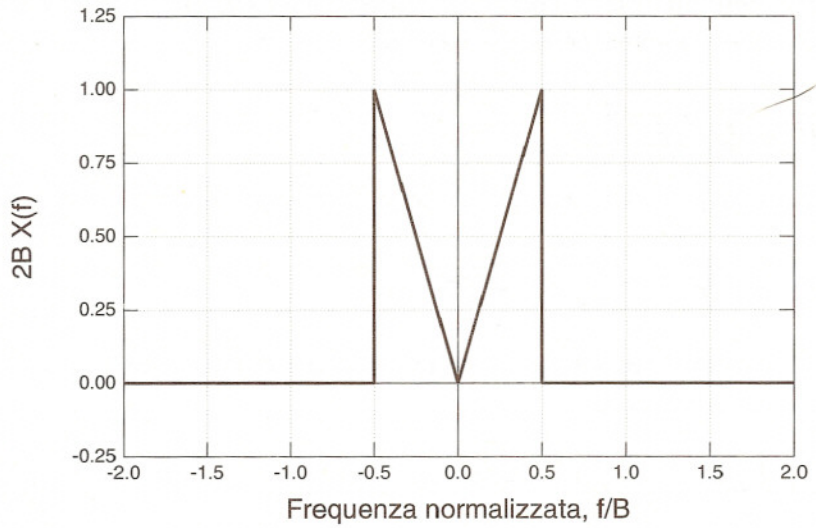


Figura 5.20 Trasformata di Fourier del segnale analogico $x(t)$

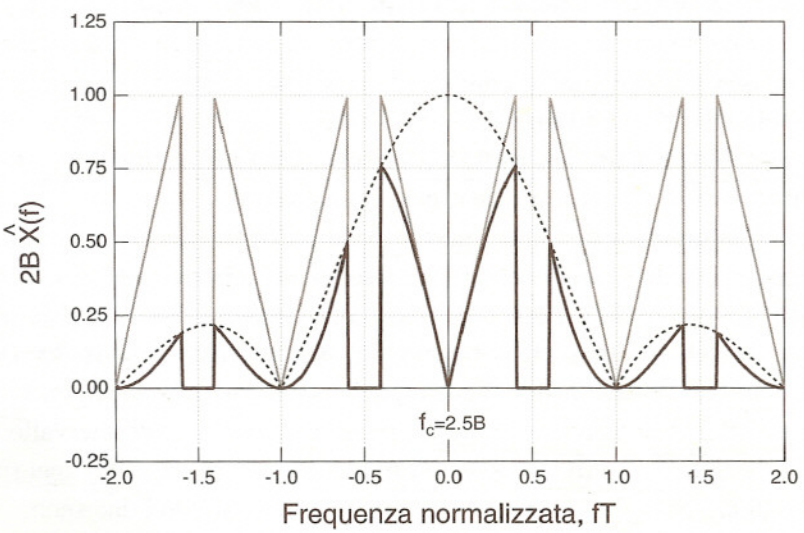


Figura 5.21 Spettro di ampiezza del segnale interpolato a mantenimento

Si può ovviare alla questione i) usando un *filtro anti-immagine* all'uscita dell'interpolatore (convertitore D/A) come indicato in Figura 5.22a. Esso è un filtro passa-basso di banda B che elimina le immagini dallo spettro del segnale interpolato, riconducendo il segnale nella banda originaria (Figura 5.22b).

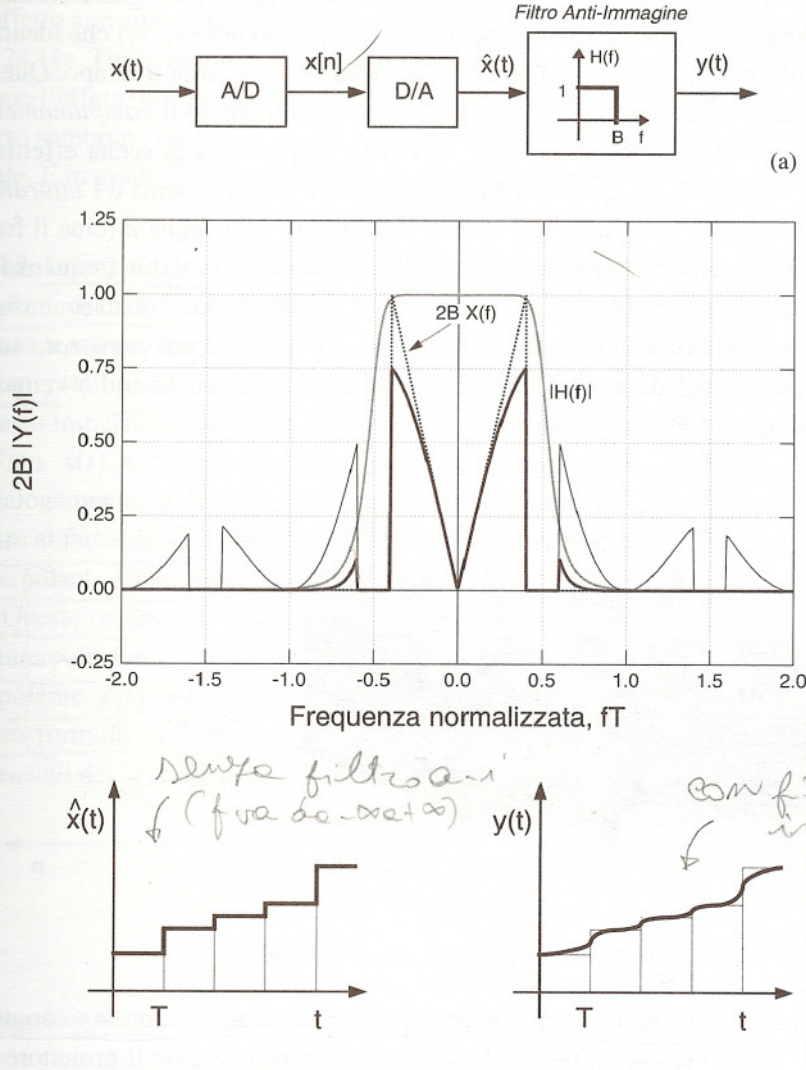


Figura 5.22 Filtro anti-immagine $H(f)$ (a) e suo effetto in frequenza (b) e nel tempo (c)

Come indicato nella Figura 5.22c, l'effetto del filtro anti-immagine in ambito temporale è quello di *smussare* il segnale costante a tratti, e quindi con disconti-

nuità, di Figura 5.19, per ricondurlo a un andamento più somigliante a quello del segnale analogico originario.

Esempio 5.8

Il segnale *cinematografico* di Figura 5.23, come abbiamo già discusso nel Paragrafo 1.2, è una *sequenza temporale di immagini*, cioè un segnale bidimensionale continuo per quel che riguarda le coordinate spaziali (x_1, x_2) che identificano il pixel nell'immagine, ma *discreto* per quel che riguarda il tempo. Questo segnale a tempo discreto $z(x_1, x_2, n]$ viene ottenuto attraverso il *campionamento* di un segnale a tempo continuo $z(x_1, x_2; t)$ che rappresenta la scena effettivamente osservata dalla cinepresa. Il *campionatore* di questo sistema è *l'otturatore* della cinepresa che fissa sulla pellicola il "campione" del segnale (cioè il fotogramma) al generico istante di scatto dell'otturatore stesso. La frequenza di campionamento è $f_c = 24$ Hz, cioè 24 fotogrammi al secondo. La sequenza di immagini ottenuta (ossia il segnale a tempo discreto) viene poi *registrata* sulla pellicola cinematografica, così come i campioni di un segnale audio vengono registrati su di un CD.

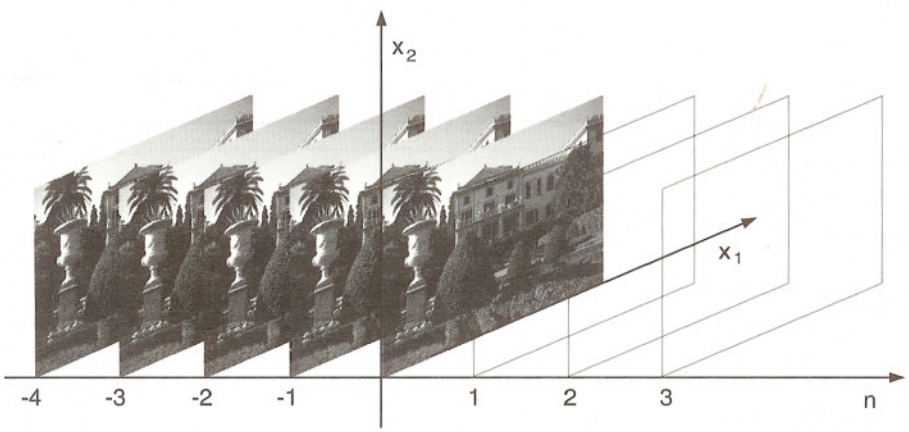


Figura 5.23 Segnale cinematografico a tempo discreto $z(x_1, x_2; n]$

INTERESSANTE !

In fase di proiezione, si desidera ricostruire il segnale *a tempo continuo* originario. Per far questo si usa un *interpolatore a mantenimento*, cioè il proiettore cinematografico. Nella proiezione (in cui l'ingrandimento sullo schermo è inessenziale perché crea una replica fedele delle immagini sulla pellicola), il valore del segnale campionato (cioè l'immagine *fissa* di ogni fotogramma) viene *man-*

tenuto per 1/24 di secondo fino all'arrivo del valore (fotogramma) successivo.

Il procedimento di interpolazione con mantenimento è efficace, cioè non si ha apparentemente percezione della "granularità" del movimento effettivamente ricostruito, perché l'occhio umano svolge la funzione di *filtro anti-immagine*. Il dato, spesso citato, di "tempo di permanenza delle immagini sulla retina" pari a circa 0.1 s porta a valutare la "banda" dell'occhio umano in circa 10 Hz, e quindi l'effetto anti-immagine filtrante è adeguato vista la frequenza di campionamento di 24 Hz. Tuttavia, nelle proiezioni cinematografiche si notano spesso *artefatti*, come l'effetto per cui le pale del rotore di un elicottero o i raggi delle ruote di un carro sembrano ruotare molto lentamente o addirittura in verso contrario a quello reale. È in grado il lettore di spiegare questi fenomeni? ❑

5.4.3 Interpolazione cardinale - Il teorema del campionamento

Le fonti di distorsione i) e ii) evidenziate nel paragrafo precedente per l'interpolatore a mantenimento possono essere attribuite alla particolare scelta dell'impulso $p(t)$ utilizzato nella formula di interpolazione (5.4.9). Le discontinuità di questo impulso inducono infatti infiniti punti di discontinuità nel segnale interpolato $\hat{x}(t)$ e causano l'allargamento illimitato della banda di $\hat{x}(t)$ stesso. Analogamente, la distorsione di ampiezza , evidenziata dalla (5.4.16), è da attribuire al fatto che nell'intervallo $[-1/2T, 1/2T]$ la trasformata $P(f)$ dell'impulso interpolante *non assume un valore costante.*

Queste osservazioni suggeriscono la possibilità di *generalizzare* l'operazione di interpolazione descritta dalla (5.4.9), scegliendo un *diverso* tipo di impulso interpolante $p(t)$ (Figura 5.24). Ovviamente, a scelte diverse di $p(t)$ corrispondono formule di interpolazione diverse, e diversi andamenti temporali e frequenziali del segnale interpolato $\hat{x}(t)$.

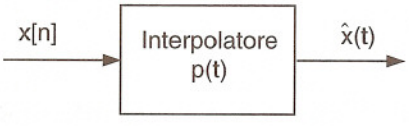


Figura 5.24 Interpolatore generalizzato

La possibilità, apparentemente banale, di generalizzare l'operazione di interpolazione assume grande importanza alla luce delle seguenti osservazioni: innanzi tutto la formula (5.4.11)

$$\hat{X}(f) = \sum_{n=-\infty}^{\infty} x[n]\, P(f)\, e^{-j2\pi nfT} = P(f) \sum_{n=-\infty}^{\infty} x[n]\, e^{-j2\pi nfT} = P(f)\, \overline{X}(f) \qquad (5.4.17)$$

è valida qualunque sia la particolare forma di $p(t)$. Se si sceglie l'impulso interpolante in modo che la sua trasformata sia *costante nell'intervallo* $[-1/2T, 1/2T]$ e nulla al di fuori, cioè

$$P(f) = T\,\mathrm{rect}(fT) \qquad (5.4.18)$$

allora si ottiene immediatamente

$$\hat{X}(f) = P(f)\,\overline{X}(f) = T\,\mathrm{rect}(fT) \cdot \frac{1}{T} \sum_{k=-\infty}^{\infty} X\left(f - \frac{k}{T}\right) = X(f) \qquad (5.4.19)$$

valida ovviamente *in assenza di aliasing*, cioè nelle ipotesi che i) $x(t)$ abbia banda limitata B, e ii) sia stata rispettata la condizione di Nyquist $f_c \geq 2B$, come mostrato nella Figura 5.25.

Questo risultato è di fondamentale importanza, ed è universalmente noto con il nome di teorema del campionamento (*sampling theorem*):

Teorema del campionamento (C. Shannon): *Un segnale il cui spettro è limitato nella banda* B *può essere ricostruito esattamente a partire dai propri campioni, purché la frequenza di campionamento non sia inferiore a* 2B.

In particolare, poiché

$$T\,\mathrm{rect}(fT) = P(f) \Leftrightarrow p(t) = \mathrm{sinc}\left(\frac{t}{T}\right) \qquad (5.4.20)$$

la formula di interpolazione risultante dalla scelta di $p(t)$ è

$$\hat{x}(t) = \sum_{n=-\infty}^{+\infty} x[n]\,\mathrm{sinc}\left(\frac{t - nT}{T}\right) \qquad (5.4.21)$$

che è nota come *formula di interpolazione cardinale*. Il nome sinc(·) assegnato a suo tempo alla funzione $\sin(\pi\alpha)/\pi\alpha$ significa infatti "seno cardinale" con riferimento all'interpolazione cardinale stessa.

La Figura 5.26 illustra un esempio di interpolazione cardinale. Il segnale analogico viene ricostruito dalla somma di una infinita serie di funzioni sinc(·), ciascuna *applicata* agli istanti nT di campionamento del segnale originario, e ciascuna *pesata* con il valore del relativo campione $x[n]$. Se *ricampioniamo* il

segnale interpolato al generico istante $t_k = kT$, per le proprietà della funzione sinc(·), solo il k-esimo fra tutti gli impulsi della sommatoria (5.4.21) produce un contributo non nullo, e pari proprio al valore $x[k] = x(kT)$ del campione del segnale di partenza (vedi la Figura 5.26).

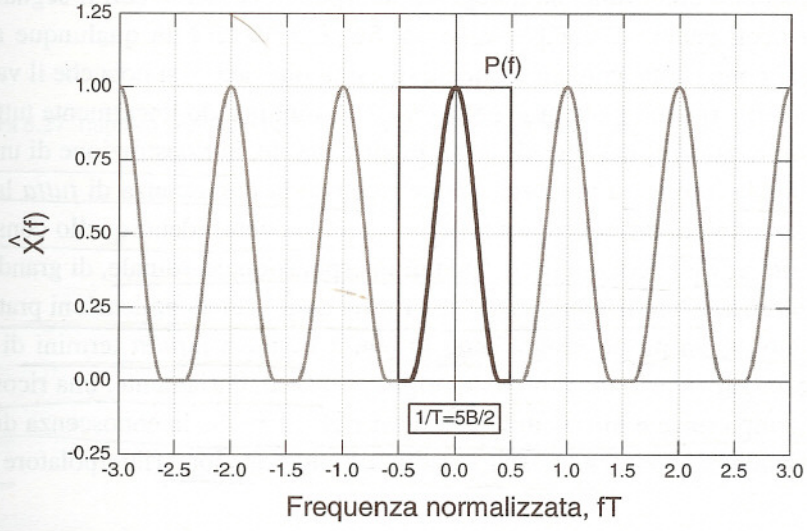


Figura 5.25 Dimostrazione grafica del teorema del campionamento

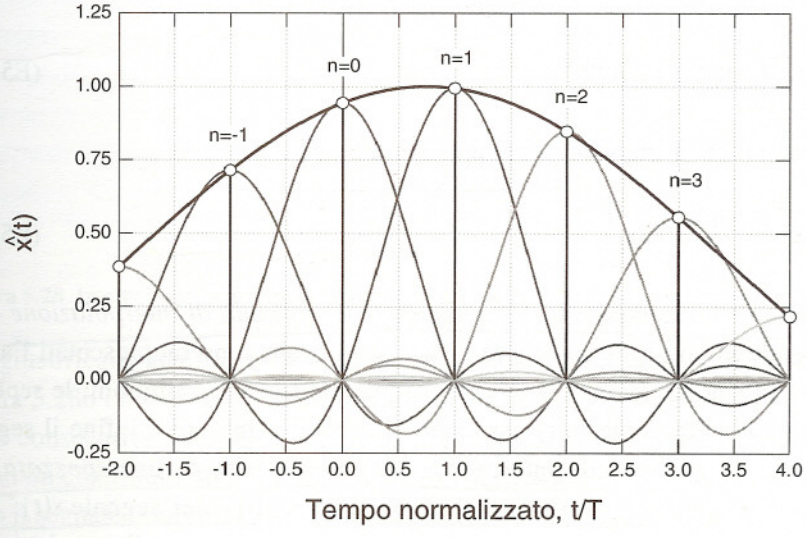


Figura 5.26 Esempio di interpolazione cardinale

Si ha infatti:

$$\hat{x}(kT) = \sum_{n=-\infty}^{+\infty} x[n]\,\mathrm{sinc}(k-n) = \sum_{n=-\infty}^{+\infty} x[n]\,\delta[k-n] = x[k] = x(kT) \qquad (5.4.22)$$

Questo risultato conferma che il segnale interpolato coincide con il segnale di partenza negli istanti di campionamento. Se si considera un qualunque altro istante non coincidente con uno di quelli di campionamento, si nota che il valore del segnale interpolato è ottenuto nella (5.4.21) combinando linearmente tutti gli *infiniti* campioni $x[n]$ del segnale $x(t)$. In altre parole, la ricostruzione di un segnale a banda limitata a un certo istante richiede la conoscenza di *tutta* la sequenza di campioni del segnale stesso, in istanti sia antecedenti quello considerato, sia *successivi*. Pertanto la formula di interpolazione cardinale, di grande rilevanza teorica, è inutilizzabile nella sua forma esatta nelle applicazioni pratiche per due motivi: in primo luogo, sono in teoria richiesti *infiniti* termini di una sommatoria per ricostruire il segnale originario; secondariamente, una ricostruzione in tempo reale è impossibile perché si richiederebbe la conoscenza di valori di segnale in istanti *successivi* a quello di interpolazione (interpolatore *non causale*).

Esempio 5.9

Riscriviamo l'espressione del segnale di uscita di un interpolatore:

$$\hat{x}(t) = \sum_{n=-\infty}^{+\infty} x[n]\,p(t-nT) \qquad (E5.9.1)$$

e supponiamo che l'impulso $p(t)$ sia (Figura 5.27)

$$p(t) = \left(1 - \frac{|t|}{T}\right)\mathrm{rect}\left(\frac{t}{2T}\right) \qquad (E5.9.2)$$

Questo impulso triangolare è caratteristico della *formula di interpolazione lineare*. Consideriamo infatti la Figura 5.28a nella quale sono rappresentati l'andamento di un segnale generico $x(t)$, la sequenza dei suoi campioni, le repliche dell'impulso interpolante $p(t)$ associate ai diversi campioni e infine il segnale interpolato $\hat{x}(t)$. Il segnale interpolato $\hat{x}(t)$ è costituito da una *spezzata* che collega i punti corrispondenti a campioni consecutivi del segnale $x(t)$: $\hat{x}(t)$ rappresenta la cosiddetta *interpolazione lineare* della sequenza di campioni.

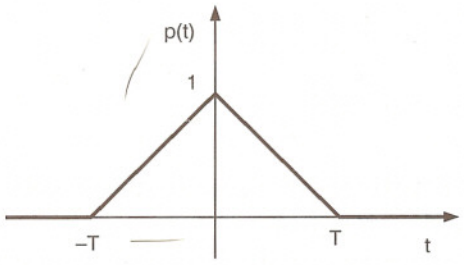
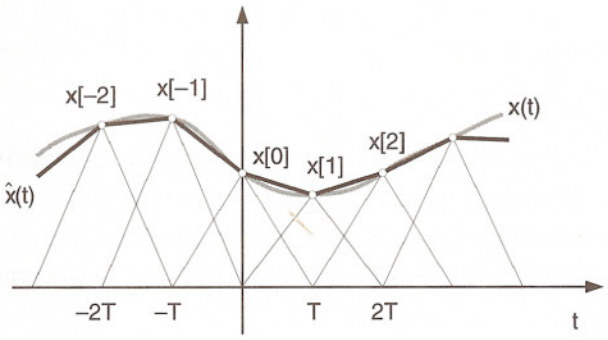


Figura 5.27 Impulso utilizzato nella formula di interpolazione lineare



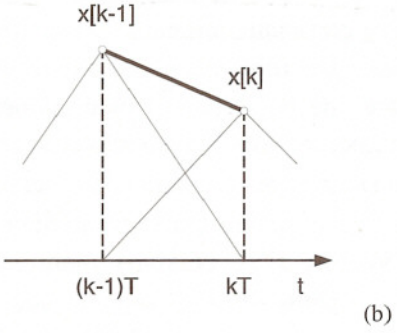


Figura 5.28 Interpolazione lineare dei campioni di un segnale $x(t)$

Per giustificare analiticamente la Figura 5.28a, osserviamo con l'aiuto della Figura 5.28b che nel generico intervallo $[(k-1)T, kT)$ compreso fra i due campioni consecutivi $x[k-1]$ e $x[k]$ solo *due* addendi della sommatoria (E5.9.1) danno un contributo non nullo, cioè quelli con $n = k-1$ e $n = k$. In questo intervallo il segnale interpolato vale allora

$$\hat{x}(t) = x[k-1]\, p(t-(k-1)T) + x[k]\, p(t-kT)$$

$$= x[k-1]\left(1-\frac{t-kT+T}{T}\right)+x[k]\left(1+\frac{t-kT}{T}\right)$$
$$= x[k]+\{x[k]-x[k-1]\}\left(\frac{t-kT}{T}\right) \qquad \text{(E5.9.3)}$$

la quale rappresenta l'equazione del segmento di retta che collega i punti corrispondenti ai campioni $x[k-1]$ e $x[k]$.

La trasformata di Fourier del segnale $\hat{x}(t)$ è data ancora dalla (5.4.11):

$$\hat{X}(f) = P(f)\,\overline{X}(f) \qquad \text{(E5.9.4)}$$

ove stavolta

$$P(f) = T\,\mathrm{sinc}^2(fT) \qquad \text{(E5.9.5)}$$

Lo spettro del segnale interpolato ha un andamento qualitativamente non dissimile da quello relativo all'interpolatore con mantenimento di Figura 5.21. Se riconsideriamo il segnale $x(t)$ il cui spettro è rappresentato in Figura 5.29a, vediamo che lo spettro di ampiezza del segnale interpolato linearmente è quello di Figura 5.29b.

Un confronto tra la Figura 5.29b e la Figura 5.21 rivela che il segnale interpolato linearmente ha uno spettro di ampiezza con immagini *più attenuate* rispetto al caso dell'interpolatore con mantenimento. Lo spettro $P(f)$ dell'impulso interpolatore decresce infatti *più rapidamente* al crescere della frequenza nel caso di interpolazione lineare che nel caso del mantenimento. Se confrontiamo il segnale generato da un interpolatore a mantenimento con quello generato da un interpolatore lineare, possiamo immediatamente osservare che il primo presenta delle discontinuità di prima specie (in corrispondenza della transizione da ciascun impulso interpolatore al successivo), mentre il secondo è un segnale *continuo*. Pertanto è da aspettarsi che il segnale con mantenimento abbia un maggior contenuto di componenti alle alte frequenze rispetto a quello prodotto da un interpolatore lineare.

Nella Figura 5.29b si nota anche che la distorsione in banda per il segnale di Figura 5.29a è abbastanza marcata. Questo deriva dalla particolare forma dello spettro del segnale di partenza, in cui sono molto ampie le componenti vicine al limite di banda B. Se lo spettro di partenza è invece più decisamente passa-basso (ossia con componenti via via digradanti con l'aumentare della frequenza), come quello di Figura 5.30a, la trasformata del segnale interpolato linearmente è come in Figura 5.30b, e la distorsione in banda è piuttosto ridotta.

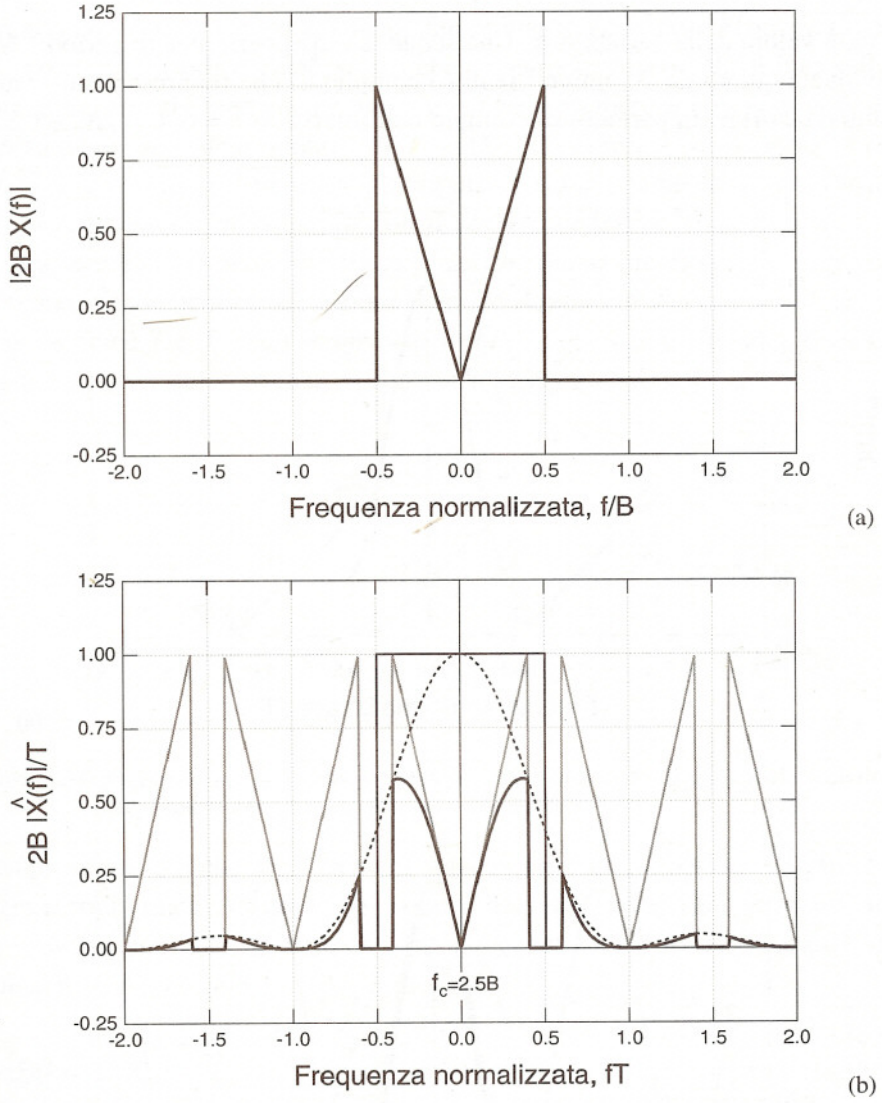


Figura 5.29 Spettro di ampiezza del segnale di partenza (a) e interpolato linearmente (b)

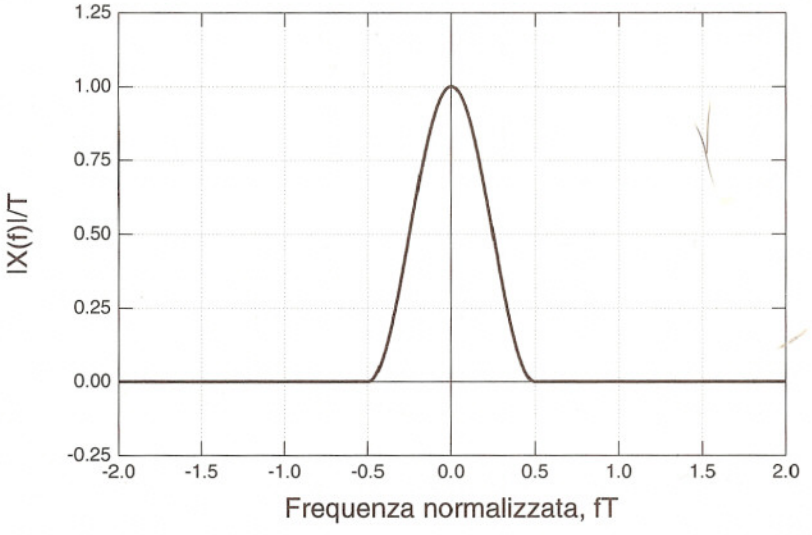
5.5 Analisi di Fourier delle sequenze periodiche

5.5.1 Trasformata discreta di Fourier

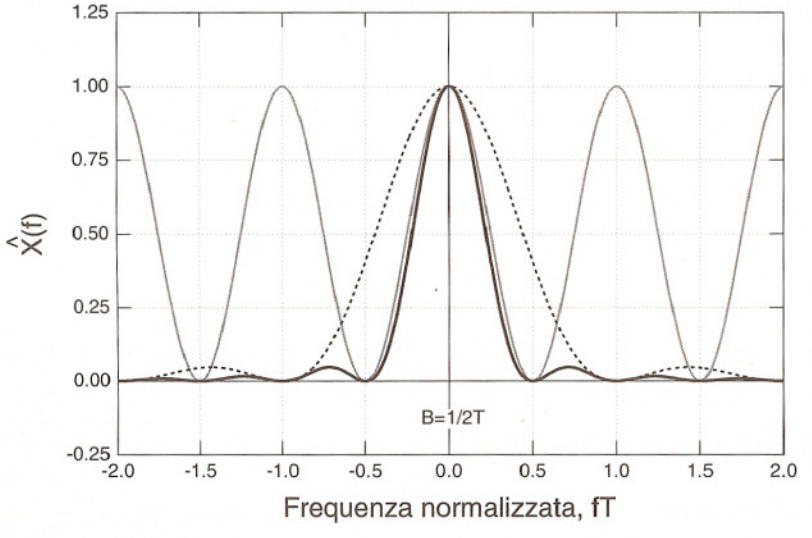
Una sequenza $x[n]$ è *periodica* se esiste un intero positivo N_0 (il *periodo* della sequenza) per il quale è verificata la seguente relazione:

$$x[n] = x[n + N_0] \tag{5.5.1}$$

per ogni valore della variabile n. Una sequenza $x[n]$ periodica di periodo N_0 è individuata quindi da N_0 numeri reali (o complessi) che rappresentano i valori assunti da $x[n]$ in un periodo, ad esempio nell'intervallo $n = 0, 1, \ldots, N_0 - 1$.



(a)



(b)

Figura 5.30 Spettro di ampiezza del segnale originario (a) e interpolato linearmente (b)

È interessante osservare che il campionamento di un segnale periodico a tempo continuo non genera necessariamente una sequenza periodica. Affinché si abbia una sequenza periodica è necessario che un numero *intero* N_0 di intervalli di campionamento sia esattamente pari a un qualche numero *intero* m di periodi di ripetizione del segnale originario: $N_0 T = mT_0$. Ciò significa che il rapporto T/T_0 deve essere *razionale*. In pratica, gli impulsi di campionamento del convertitore A/D devono essere *sincronizzati* con il segnale periodico analogico: essi non possono avere una cadenza arbitraria senza un legame preciso con la cadenza di ripetizione fondamentale del segnale dato. Se il rapporto T/T_0 *non è* un numero razionale (cioè T e T_0 sono *incommensurabili*), l'operazione di campionamento *non* dà origine a una sequenza $x[n]$ periodica (Figura 5.31).

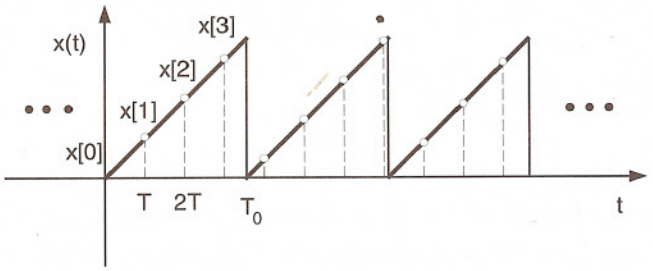


Figura 5.31 Sequenza non periodica generata dal campionamento di un segnale periodico a tempo continuo

Supponiamo ora che $x[n]$ sia una sequenza periodica di periodo N_0. Essa può essere rappresentata mediante uno sviluppo del tutto analogo alla serie di Fourier per i segnali periodici a tempo continuo, chiamato *serie discreta* o *antitrasformata discreta di Fourier*:

$$x[n] = \sum_{k=0}^{N_0-1} \overline{X}_k \, e^{j\frac{2\pi kn}{N_0}} \tag{5.5.2}$$

La *sequenza* $\overline{X}_k$ dei coefficienti discreti di Fourier è comunemente chiamata *trasformata discreta di Fourier* della sequenza periodica data ed è pari a

$$\overline{X}_k = \frac{1}{N_0} \sum_{n=0}^{N_0-1} x[n] \, e^{-j\frac{2\pi kn}{N_0}} \tag{5.5.3}$$

Notiamo le analogie tra le (5.5.2-5.5.3) e le corrispondenti relazioni di sintesi e analisi per i segnali periodici a tempo continuo:

$$\boxed{\begin{aligned} x[n] &= \sum_{k=0}^{N_0-1} \overline{X}_k \, e^{j\frac{2\pi kn}{N_0}} \\ \overline{X}_k &= \frac{1}{N_0} \sum_{n=0}^{N_0-1} x[n] \, e^{-j\frac{2\pi kn}{N_0}} \end{aligned}} \qquad \boxed{\begin{aligned} x(t) &= \sum_{k=-\infty}^{\infty} X_k \, e^{j\frac{2\pi kt}{T_0}} \\ X_k &= \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t) \, e^{-j\frac{2\pi kt}{T_0}} dt \end{aligned}} \tag{5.5.4}$$

Per i segnali periodici a tempo continuo la rappresentazione mediante serie di Fourier comporta una somma *infinita* di termini; nel caso di sequenze periodiche, invece, la rappresentazione mediante antitrasformata discreta consiste in una somma con un numero *finito* di addendi. Infatti la trasformata di una sequenza $x[n]$ periodica di periodo N_0 è essa stessa *periodica* con il medesimo periodo:

$$\overline{X}_{k+N_0} = \frac{1}{N_0} \sum_{n=0}^{N_0-1} x[n] \, e^{-j\frac{2\pi(k+N_0)n}{N_0}} = \frac{1}{N_0} \sum_{n=0}^{N_0-1} x[n] \, e^{-j\frac{2\pi kn}{N_0}} \, e^{-j2\pi n} =$$

$$= \frac{1}{N_0} \sum_{n=0}^{N_0-1} x[n] \, e^{-j\frac{2\pi kn}{N_0}} = \overline{X}_k \tag{5.5.5}$$

La sequenza periodica è espressa, come nel caso del segnale periodico a tempo continuo, da una somma di oscillazioni sinusoidali a frequenze *armoniche*, cioè *multiple di una frequenza fondamentale*. Riscriviamo infatti l'equazione di sintesi come segue:

$$x[n] = \sum_{k=0}^{N_0-1} \overline{X}_k \, e^{j2\pi\frac{k}{N_0 T}nT} \tag{5.5.6}$$

Si nota che i vari esponenziali complessi nella scomposizione oscillano alle frequenze $f_k = k/(N_0 T)$, $k = 0,...,N_0 - 1$, che a buon diritto possono chiamarsi le *armoniche* relative al periodo di ripetizione N_0, ovvero alla frequenza fondamentale $1/(N_0 T)$. Una piccola diversità tra le equazioni di analisi a tempo continuo e discreto nella (5.5.4) sta nel fatto che l'integrale per ricavare il coefficiente di Fourier si calcola sull'intervallo *simmetrico* $(-T_0/2, T_0/2)$, mentre la trasformata discreta viene calcolata su di un intervallo *asimmetrico* $[0, N_0 - 1]$. La spiegazione è semplice: quando N_0 è un numero *dispari*, è semplice calcolare la trasformata discreta anche sull'intervallo simmetrico $[-(N_0 - 1)/2, (N_0 - 1)/2]$, ma se N_0 è un numero *pari*, non è possibile trovare un intervallo simmetrico di ampiezza pari al periodo N_0. Si preferisce quindi unificare i due casi usando l'intervallo asimmetrico *destro* $[0, N_0 - 1]$.

Dimostriamo adesso che dalla relazione (5.5.2) di *sintesi* (antitrasformata

discreta) discende la (5.5.3) di analisi (trasformata discreta). Moltiplichiamo ambo i membri della (5.5.2) per il fattore $e^{-j2\pi nm/N_0}$ $(0 \le m \le N_0 - 1)$ ed effettuiamo l'operazione di somma sul periodo:

$$\sum_{n=0}^{N_0-1} x[n]\, e^{-j\frac{2\pi nm}{N_0}} = \sum_{n=0}^{N_0-1} \sum_{k=0}^{N_0-1} \overline{X}_k\, e^{-j\frac{2\pi nm}{N_0}} e^{j\frac{2\pi nk}{N_0}} \tag{5.5.7}$$

Sviluppando il secondo membro della (5.5.7) si ricava

$$\sum_{n=0}^{N_0-1} \sum_{k=0}^{N_0-1} \overline{X}_k\, e^{-j\frac{2\pi nm}{N_0}} e^{j\frac{2\pi nk}{N_0}} = \sum_{k=0}^{N_0-1} \overline{X}_k \sum_{n=0}^{N_0-1} e^{j\frac{2\pi n(k-m)}{N_0}} \tag{5.5.8}$$

La seconda sommatoria a secondo membro della (5.5.8) diviene

$$\sum_{n=0}^{N_0-1} e^{j\frac{2\pi n(k-m)}{N_0}} = \sum_{n=0}^{N_0-1} \left(e^{j\frac{2\pi(k-m)}{N_0}} \right)^n = \frac{1 - e^{j2\pi(k-m)}}{1 - e^{j\frac{2\pi(k-m)}{N_0}}} = 0 \quad , \quad k \neq m \tag{5.5.9}$$

mentre, per $k = m$, si ha

$$\sum_{n=0}^{N_0-1} e^{j\frac{2\pi n(k-m)}{N_0}} = \sum_{n=0}^{N_0-1} 1 = N_0 \tag{5.5.10}$$

Pertanto dalla (5.5.8) si trova che

$$\sum_{n=0}^{N_0-1} \sum_{k=0}^{N_0-1} \overline{X}_k\, e^{-j\frac{2\pi nm}{N_0}} e^{j\frac{2\pi nk}{N_0}} = \sum_{k=0}^{N_0-1} \overline{X}_k\, \delta[k-m]\, N_0 = N_0 \overline{X}_m \tag{5.5.11}$$

Sostituendo la (5.5.11) nella (5.5.7) si ricava infine

$$\sum_{n=0}^{N_0-1} x[n]\, e^{-j\frac{2\pi nm}{N_0}} = N_0 \overline{X}_m \tag{5.5.12}$$

da cui segue immediatamente l'equazione di analisi (5.5.3).

Le proprietà della trasformata discreta sono molto simili a quelle già discusse nel Paragrafo 3.2 a proposito della trasformata dei segnali analogici, e non verranno ulteriormente ribadite.

La maggior parte dei teoremi riguardanti la trasformata discreta possono inoltre essere facilmente adattati o ricavati sulla base dei teoremi descritti nel Paragrafo 5.3. Ci limitiamo a esaminare in dettaglio il caso dei teoremi del *prodotto* e della *convoluzione* per le sequenze periodiche.

Esempio 5.10

Calcoliamo la trasformata discreta della sequenza

$$x[n] = \cos\left(\frac{n\pi}{4}\right) \tag{E5.10.1}$$

che è evidentemente periodica di periodo $N_0 = 8$. Dalle formule di Eulero,

$$x[n] = \left(e^{j\pi n/4} + e^{-j\pi n/4}\right) = \frac{1}{2}e^{j2\pi n/8} + \frac{1}{2}e^{-j2\pi n/8}$$
$$= \frac{1}{2}e^{j2\pi n/8} + \frac{1}{2}e^{j[2\pi n - 2\pi n/8]} = \frac{1}{2}e^{j2\pi n/8} + \frac{1}{2}e^{j7\cdot 2\pi n/8} \tag{E5.10.2}$$

per cui, senza bisogno di effettuare materialmente la trasformata, si può concludere direttamente che

$$\overline{X}_1 = \frac{1}{2}\,, \quad \overline{X}_7 = \frac{1}{2} \tag{E5.10.3a}$$

e

$$\overline{X}_k = 0 \;\;, \quad k = 0,2,3,4,5,6 \tag{E5.10.3b}$$

Questo risultato ricorda quello ricavato nell'Esempio 2.1 per il calcolo dei coefficienti di Fourier del segnale $x(t) = \cos(2\pi f_0 t)$. Il segnale $x[n]$ è un'oscillazione cosinusoidale (discreta) con frequenza $f_0 = 1/(8T)$, e quindi le uniche armoniche diverse da 0 sono quelle con frequenza $\pm f_0$, cioè con indice di armonica pari a ± 1, in quanto f_0 coincide con la frequenza *fondamentale* $1/(N_0 T)$. La (E5.10.3a) non è in contrasto con quest'osservazione, visto che

$$\overline{X}_{-1} = \overline{X}_7 = \frac{1}{2} \tag{E5.10.4}$$

per la proprietà di *periodicità* della trasformata discreta. ❑

Esempio 5.11

Calcoliamo la trasformata discreta della sequenza $x[n]$ periodica di periodo $N_0 = 8$, definita sul periodo-base come segue:

$$x[n] = \begin{cases} 1 & 0 \le n \le 3 \\ 0 & 4 \le n \le 7 \end{cases} \tag{E5.11.1}$$

Questa sequenza è in pratica un *treno di impulsi rettangolari* a tempo discreto, come è chiaro dalla rappresentazione in Figura 5.32a.

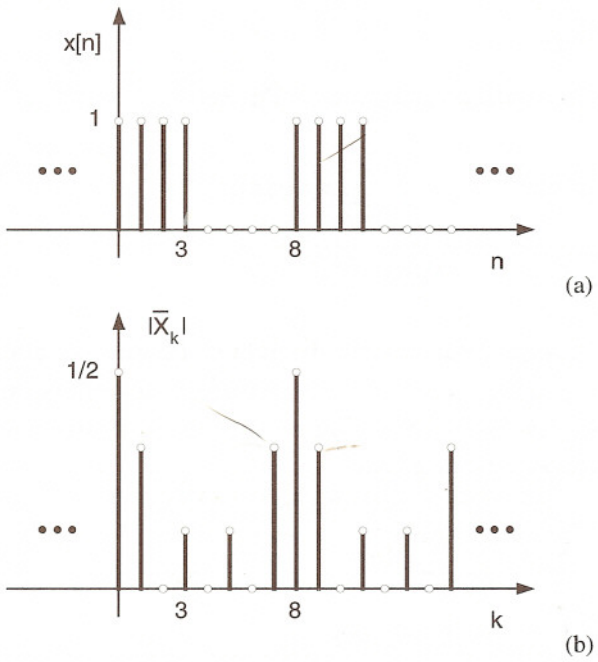


Figura 5.32 Segnale periodico dell'Esempio 5.11 (a) e suo spettro di ampiezza $|\overline{X}_k|$ (b)

La trasformata discreta è

$$\overline{X}_k = \frac{1}{N_0}\sum_{n=0}^{N_0-1} x[n]\, e^{-j\frac{2\pi kn}{N_0}} = \frac{1}{8}\sum_{n=0}^{3} e^{-j\frac{2\pi kn}{8}} = \frac{1}{8}\sum_{n=0}^{3}\left(e^{-j\frac{\pi k}{4}}\right)^n \qquad \text{(E5.11.2)}$$

Si può applicare la formula della somma di una progressione geometrica quando $k \neq 0$, mentre quando $k = 0$ si ha banalmente $\overline{X}_0 = 4/8 = 1/2$. Riassumendo queste osservazioni si ha:

$$\overline{X}_k = \begin{cases} 1/2 & k = 0 \\ \left(\frac{1}{8}\right)\dfrac{1-e^{-jk\pi}}{1-e^{-j\pi k/4}} & k \neq 0 \end{cases} \qquad \text{(E5.11.3)}$$

Si trova poi che $\overline{X}_2 = \overline{X}_4 = \overline{X}_6 = 0$, che $|\overline{X}_1| = |\overline{X}_7| = 1/[8\sin(\pi/8)] \cong 0.327$, e che inoltre $|\overline{X}_3| = |\overline{X}_5| = 1/[8\sin(3\pi/8)] \cong 0.135$. Lo spettro di ampiezza del segnale dato (cioè il grafico di $|\overline{X}_k|$) è dunque quello di Figura 5.32b. ❑

5.5.2 Teorema del *prodotto*

Consideriamo adesso la sequenza (periodica) $p[n]$ data dal prodotto fra la sequenza $x[n]$ e la sequenza $y[n]$ entrambe periodiche di periodo N_0

$$p[n] = x[n]\, y[n] \tag{5.5.13}$$

e calcoliamone la trasformata discreta di Fourier:

$$\overline{P}_k = \frac{1}{N_0}\sum_{n=0}^{N_0-1} p[n]\, e^{-j2\pi\frac{nk}{N_0}} = \frac{1}{N_0}\sum_{n=0}^{N_0-1} x[n]y[n]\, e^{-j2\pi\frac{nk}{N_0}}$$
$$= \frac{1}{N_0}\sum_{n=0}^{N_0-1}\sum_{m=0}^{N_0-1} \overline{X}_m\, e^{j2\pi\frac{nm}{N_0}} \cdot y[n]\, e^{-j2\pi\frac{nk}{N_0}} \tag{5.5.14}$$

ove $x[n]$ è stata scomposta in serie discreta di Fourier. In questo passaggio, è stata usata una variabile "muta" m nell'operazione di antitrasformazione per non creare ambiguità con la variabile k da cui dipende la trasformata $\overline{P}$. Invertendo l'ordine delle sommatorie si ricava

$$\overline{P}_k = N_0 \cdot \frac{1}{N_0}\sum_{m=0}^{N_0-1} \overline{X}_m \frac{1}{N_0}\sum_{n=0}^{N_0-1} y[n] e^{-j2\pi\frac{n(k-m)}{N_0}}$$
$$= N_0 \cdot \frac{1}{N_0}\sum_{m=0}^{N_0-1} \overline{X}_m \overline{Y}_{k-m} = N_0 \cdot \overline{X}_k \otimes \overline{Y}_k \tag{5.5.15}$$

ove naturalmente la convoluzione tra le trasformate discrete è una somma di convoluzione *ciclica* tra le due sequenze periodiche $\overline{X}_k$ e $\overline{Y}_k$ in ambito frequenziale. In conclusione:

$$p[n] = x[n]\, y[n] \Leftrightarrow \overline{P}_k = \sum_{m=0}^{N_0-1} \overline{X}_m \overline{Y}_{k-m} = N_0 \cdot \overline{X}_k \otimes \overline{Y}_k \tag{5.5.16}$$

5.5.3 Teorema della *convoluzione*

Consideriamo ora la sequenza $z[n]$ come *somma di convoluzione ciclica* tra le due sequenze $x[n]$ e $y[n]$, entrambe periodiche di periodo N_0:

$$z[n] = x[n] \otimes y[n] = \frac{1}{N_0}\sum_{m=0}^{N_0-1} x[m]\, y[n-m] = \frac{1}{N_0}\sum_{m=0}^{N_0-1} y[m]\, x[n-m] \tag{5.5.17}$$

Questa somma di convoluzione gode ovviamente delle stesse proprietà *commutativa*, *associativa* e *distributiva* già citate per la somma di convoluzione tra sequenze aperiodiche.

Calcoliamo poi la trasformata discreta di $z[n]$:

$$\overline{Z}_k = \frac{1}{N_0}\sum_{n=0}^{N_0-1} z[n]\, e^{-j2\pi\frac{nk}{N_0}} = \frac{1}{N_0}\sum_{n=0}^{N_0-1}\frac{1}{N_0}\sum_{m=0}^{N_0-1} x[m]\, y[n-m]\, e^{-j2\pi\frac{nk}{N_0}}$$

$$= \frac{1}{N_0}\sum_{m=0}^{N_0-1} x[m]\frac{1}{N_0}\sum_{n=0}^{N_0-1} y[n-m]e^{-j2\pi\frac{nk}{N_0}}$$

$$= \frac{1}{N_0}\sum_{m=0}^{N_0-1} x[m]\overline{Y}_k\, e^{-j2\pi\frac{mk}{N_0}} = \overline{X}_k \cdot \overline{Y}_k \qquad (5.5.18)$$

In conclusione, possiamo enunciare il teorema della convoluzione (ciclica) nella forma

$$x[n] \otimes y[n] \Leftrightarrow \overline{X}_k \cdot \overline{Y}_k \qquad (5.5.19)$$

Esempio 5.12

Calcoliamo la somma di convoluzione $z[n]$ tra le due sequenze

$$x[n] = u[n] - u[n-4] \quad , \quad y[n] = \left(\frac{1}{2}\right)^n u[n] \qquad (E5.12.1)$$

La prima delle due è un impulso rettangolare di durata 4, mentre la seconda è una sequenza esponenziale unilatera. La convoluzione che cerchiamo è data da

$$z[n] = x[n] \otimes y[n] = \sum_{m=-\infty}^{+\infty} x[n-m]y[m] \qquad (E5.12.2)$$

che svolgeremo graficamente. La Figura 5.33a, che rappresenta i segnali $x[-m]$ e $y[m]$ coinvolti nel calcolo, suggerisce che il risultato della convoluzione è differente a seconda che $n < 0$, $n \geq 3$, o $0 \leq n \leq 2$ (ricordiamo che n è la traslazione che si deve dare al segnale $x[-m]$ nel calcolo della somma (E5.12.2)). Nel primo caso, non c'è "sovrapposizione" tra i due segnali, il loro prodotto è nullo e la convoluzione è pure nulla. Nel secondo caso, quattro campioni del segnale $y[m]$ contribuiscono al risultato, che è pari a

$$z[n] = \sum_{m=n-3}^{n} y[m] = \left(\frac{1}{2}\right)^{n-3} + \left(\frac{1}{2}\right)^{n-2} + \left(\frac{1}{2}\right)^{n-1} + \left(\frac{1}{2}\right)^{n} = \frac{15}{8}\left(\frac{1}{2}\right)^{n-3} \qquad (E5.12.3)$$

Nel terzo caso, invece, i campioni che contribuiscono al risultato sono quelli con $0 \leq m \leq n$, e quindi il risultato è

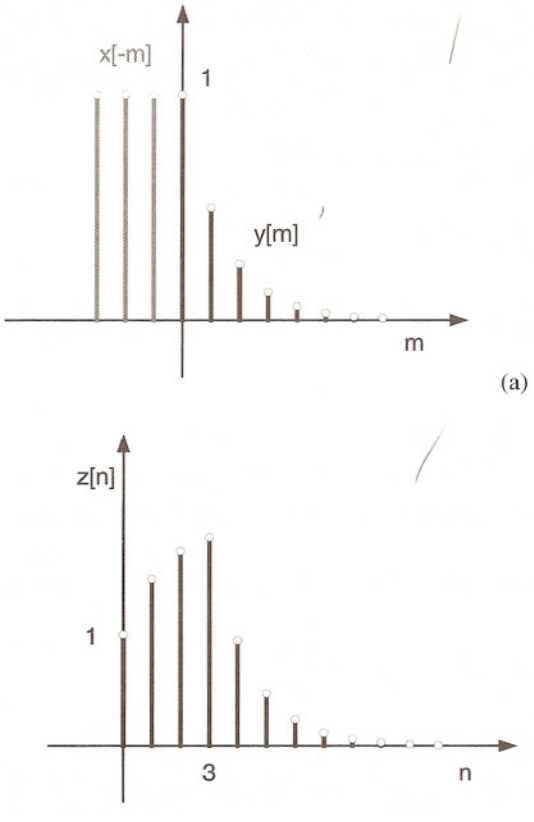


Figura 5.33 Calcolo (a) e risultato (b) della convoluzione tra le sequenze dell'Esempio 5.12

$$z[n]=\sum_{m=0}^{n} y[m]=\left(\frac{1}{2}\right)^{0}+\cdots+\left(\frac{1}{2}\right)^{n}=2-\left(\frac{1}{2}\right)^{n} \qquad \text{(E5.12.4)}$$

Riassumendo:

$$z[n]=\begin{cases}0 & n<0\\ 2-\left(\frac{1}{2}\right)^{n} & 0\le n\le 2\\ \frac{15}{8}\left(\frac{1}{2}\right)^{n-3} & n\ge 3\end{cases} \qquad \text{(E5.12.5)}$$

come rappresentato nella Figura 5.33b.

❑

5.5.4 Periodicizzazione di una sequenza aperiodica

Nel Paragrafo 3.5 abbiamo studiato la periodicizzazione di periodo T_0 di un segnale aperiodico a tempo continuo $x(t)$ per ottenere un segnale $y(t)$ periodico

$$y(t) = \sum_{m=-\infty}^{+\infty} x(t - mT_0) \tag{5.5.20}$$

e abbiamo ricavato la relazione di *campionamento in frequenza* fra i coefficienti Y_k dello sviluppo in serie di Fourier di $y(t)$ e la trasformata continua di Fourier $X(f)$ di $x(t)$:

$$Y_k = \frac{1}{T_0} X\left(\frac{k}{T_0}\right) \tag{5.5.21}$$

Vogliamo ora ricavare la relazione analoga per i segnali a tempo discreto. Costruiamo dunque la sequenza $y[n]$ periodica di periodo N_0 a partire dalla sequenza aperiodica $x[n]$:

$$y[n] = \sum_{m=-\infty}^{+\infty} x[n - mN_0] \tag{5.5.22}$$

La trasformata discreta (cioè il k-esimo coefficiente della serie discreta di Fourier) di $y[n]$ è

$$\overline{Y}_k = \frac{1}{N_0} \sum_{n=0}^{N_0-1} y[n]\, e^{-j\frac{2\pi nk}{N_0}} \tag{5.5.23}$$

Sostituendo l'espressione della sequenza periodicizzata si ottiene

$$\overline{Y}_k = \frac{1}{N_0} \sum_{n=0}^{N_0-1} \sum_{m=-\infty}^{+\infty} x[n - mN_0]\, e^{-j\frac{2\pi nk}{N_0}} = \frac{1}{N_0} \sum_{m=-\infty}^{\infty} \sum_{n=0}^{N_0-1} x[n - mN_0]\, e^{-j\frac{2\pi nk}{N_0}} \tag{5.5.24}$$

Se nella relazione precedente si effettua il cambiamento di variabile $p = n - mN_0$ si ha

$$\begin{aligned} \overline{Y}_k &= \frac{1}{N_0} \sum_{m=-\infty}^{+\infty} \sum_{p=-mN_0}^{-N_0(m-1)-1} x[p]\, e^{-j\frac{2\pi(p+mN_0)k}{N_0}} = \frac{1}{N_0} \sum_{m=-\infty}^{+\infty} \sum_{p=-mN_0}^{-(m-1)N_0-1} x[p]\, e^{-j\frac{2\pi pk}{N_0}} e^{-j2\pi mk} \\ &= \frac{1}{N_0} \sum_{m=-\infty}^{+\infty} \sum_{p=-mN_0}^{-(m-1)N_0-1} x[p]\, e^{-j\frac{2\pi pk}{N_0}} \end{aligned} \tag{5.5.25}$$

La sequenza "sommanda" a secondo membro nella (5.5.25) non dipende dall'indice della serie m. Tale indice agisce infatti solo sugli *estremi di somma* della

sommatoria interna. Ci si rende allora conto facilmente, con l' ausilio della Figura 5.34, che, al variare di m tra $-\infty$ e $+\infty$, gli intervalli di somma $[-mN_0, -(m-1)N_0 - 1]$ *della stessa sequenza sommanda* ricoprono tutti i numeri interi relativi senza sovrapposizioni. Se ne conclude che la doppia sommatoria della (5.5.25) può essere riscritta come *un'unica* sommatoria, cioè

$$\overline{Y}_k = \frac{1}{N_0} \sum_{p=-\infty}^{+\infty} x[p]\, e^{-j2\pi p\left(\frac{k}{N_0 T}\right)T} \tag{5.5.26}$$

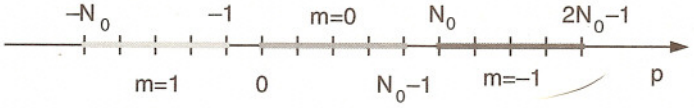


Figura 5.34 Spiegazione grafica della formula (5.5.26)

Infine, richiamando la definizione di trasformata di una sequenza aperiodica

$$\overline{X}(f) = \sum_{n=-\infty}^{+\infty} x[n]\, e^{-j2\pi nfT} \tag{5.5.27}$$

si vede che la (5.5.26) può essere riscritta come

$$\overline{Y}_k = \frac{1}{N_0} \overline{X}(f)\Big|_{f=\frac{k}{N_0 T}} = \frac{1}{N_0} \overline{X}\left(\frac{k}{N_0 T}\right) \tag{5.5.28}$$

che rappresenta la relazione cercata di "campionamento in frequenza".

Esempio 5.13

Ricaviamo la trasformata discreta di Fourier del *treno di impulsi rettangolari discreto* $y[n]$ rappresentato in Figura 5.35.

Il segnale periodico $y[n]$ si può ottenere dalla *periodicizzazione* con periodo N_0 dell'impulso rettangolare aperiodico $x[n]$ dell'Esempio 5.2, per il quale si ha

$$\overline{X}(f) = e^{-j\pi(N-1)fT} \frac{\sin(N\pi fT)}{\sin(\pi fT)} \tag{E5.13.1}$$

Considerando ora la relazione (5.5.28) di campionamento in frequenza, si ricava immediatamente la trasformata discreta di $y[n]$:

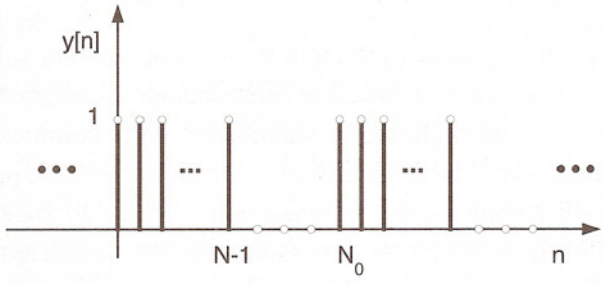


Figura 5.35 Treno di impulsi rettangolari a tempo discreto

$$\overline{Y}_k = \frac{1}{N_0}\overline{X}\left(\frac{k}{N_0 T}\right) = e^{-j\pi k(N-1)/N_0}\,\frac{\sin(\pi kN/N_0)}{N_0\sin(\pi k/N_0)} \qquad \text{(E5.13.2)}$$

Se $N_0 = 8$ e $N = 4$ si riottiene evidentemente il risultato dell'Esempio 5.11. ❑

5.6 Cenno agli algoritmi veloci di trasformata discreta (FFT)

5.6.1 Complessità di calcolo della trasformata discreta

Supponiamo di avere disponibile nella memoria di un calcolatore gli N_0 valori-base di una sequenza periodica $x[n]$ e di volerne calcolare numericamente la trasformata discreta

$$\overline{X}_k = \frac{1}{N_0}\sum_{n=0}^{N_0-1} x[n]\,e^{-j\frac{2\pi kn}{N_0}}$$
$$= \frac{1}{N_0}\left\{x[0]\cdot e^{-j0} + x[1]\cdot e^{-j\frac{2\pi k}{N_0}} + x[2]\cdot e^{-j\frac{2\pi 2k}{N_0}} + \ldots + x[N_0-1]\cdot e^{-j\frac{2\pi(N_0-1)k}{N_0}}\right\}$$

(5.6.1*a*)

Cerchiamo allora di valutare l'impegno di calcolo in termini di numero di operazioni che il calcolatore deve eseguire per ricavare gli N_0 valori della trasformata. Notiamo in via preliminare che le considerazioni che faremo a proposito del calcolo di una trasformata sono validi anche per una *antitrasformata*, purché si faccia l'ipotesi che i valori $x[0]\ldots x[N_0-1]$ della sequenza siano *complessi* e si ignori il fattore di scala $1/N_0$. Se infatti riscriviamo l'antitrasformata discreta di Fourier nella forma della (5.6.1*a*) troviamo:

$$x[n] = \overline{X}_0 \cdot e^{j0} + \overline{X}_1 \cdot e^{j\frac{2\pi n}{N_0}} + \overline{X}_2 \cdot e^{j\frac{2\pi 2n}{N_0}} + \ldots + \overline{X}_{N_0-1} \cdot e^{j\frac{2\pi(N_0-1)n}{N_0}} \qquad (5.6.1b)$$

Si devono cioè effettuare sostanzialmente le stesse operazioni sulle varie quantità $W_{N_0}^{nk} \triangleq e^{j2\pi nk/N_0}$, salvo un inessenziale cambiamento di segno dell'esponente. In questo modo, potremo unificare la valutazione della complessità dell'algoritmo di trasformata diretta e di quello di *antitrasformata*. Supponendo che i fattori esponenziali complessi che figurano nelle (5.6.1*a-b*) siano *precalcolati,* cioè già disponibili in memoria, la determinazione del coefficiente $\overline{X}_k$ (ovvero del campione $x[n]$) richiede N_0 moltiplicazioni complesse ed $N_0 - 1$ addizioni complesse. Tenendo conto che un'addizione complessa richiede in realtà 2 addizioni reali, e una moltiplicazione complessa richiede 4 moltiplicazioni reali e 2 addizioni, sono necessarie complessivamente $8N_0 - 2$ operazioni (reali) per ogni valore di k. Poiché k assume tutti i valori compresi tra 0 e $N_0 - 1$, il numero complessivo di operazioni da compiere per calcolare la trasformata discreta di Fourier (TDF) di una sequenza periodica di periodo N_0 è pari a

$$N_{TDF}(N_0) = (8N_0 - 2)\, N_0 = 8N_0^2 - 2N_0 \cong 8N_0^2 \qquad (5.6.2)$$

ove l'ultima approssimazione è stata fatta supponendo che il parametro N_0 assuma un valore grande[1]. Possiamo notare che la *complessità di calcolo* (o *computazionale*) della trasformata discreta è di tipo *quadratico* nell'ordine N_0 di trasformazione. Se immaginiamo di disporre di un elaboratore che svolge le operazioni con una cadenza f_{clock} pari a 100 MHz (cioè 100 milioni di operazioni al secondo), e prendiamo per N_0 il valore tipico $N_0 = 2^{10} = 1024$, il tempo necessario per effettuare il calcolo suddetto è pari a

$$T_{N_0} = N_{TOT} / f_{clock} \cong 8 \cdot 2^{20} \cdot 10^{-8} \cong 8 \cdot 10^6 \cdot 10^{-8} = 80 \text{ ms} \qquad (5.6.3)$$

Supponiamo adesso che la sequenza $x[n]$ venga generata *campionando* un segnale analogico $x(t)$ a frequenza f_c, e che si vogliano calcolare trasformate discrete *ripetute* su spezzoni temporali adiacenti (le cosiddette *finestre)* di N_0 campioni consecutivi della sequenza. Se in un tempo pari a 80 ms riusciamo a elaborare 1024 campioni, allora per poter operare in *tempo reale* si deve utilizzare una frequenza di campionamento f_c tale che

[1] In questa ipotesi, è possibile anche trascurare le eventuali $2N_0$ moltiplicazioni finali per il fattore di scala $1/N_0$ nell'algoritmo di trasformazione diretta.

$$f_c \le \frac{N_0}{T_{N_0}} \frac{1024}{80 \cdot 10^{-3}} = 12.8 \text{ kHz} \tag{5.6.4}$$

altrimenti nuovi campioni di una nuova finestra di segnale vengono presi *prima* che il precedente calcolo della trasformata sia terminato. Questo risultato rappresenta un vincolo sulla banda B del segnale $x(t)$ da elaborare poiché la condizione di Nyquist richiede che valga $B \le f_c/2$.

Un deciso miglioramento della velocità di elaborazione può essere conseguito utilizzando un *algoritmo*[2] *veloce di calcolo* della trasformata discreta che, sfruttando particolari simmetrie insite nei fattori $W_{N_0}^{nk}$ della trasformata stessa, riduce la *complessità computazionale* del medesimo. Tale algoritmo, noto come *Fast Fourier Transform* (FFT)[3], a parità di frequenza di clock dell'elaboratore, permette l'utilizzo di frequenze di campionamento notevolmente maggiori rispetto a quella indicata dalla relazione (5.6.4). Cerchiamo quindi di dare un'idea del principio di funzionamento della FFT e di valutarne il grado di complessità.

Il più semplice algoritmo di FFT si applica quando l'*ordine* N_0 della trasformata è una potenza di 2, cioè $N_0 = 2^M$. Riscriviamo in tal caso la (5.6.1) ignorando il fattore di scala $1/N_0$ e suddividendo gli addendi in due gruppi:

$$\begin{aligned} \overline{X}_k &= \sum_{m=0}^{N_0/2-1} x[2m]\, e^{-j\frac{2\pi(2m)k}{N_0}} + \sum_{m=0}^{N_0/2-1} x[2m+1]\, e^{-j\frac{2\pi(2m+1)k}{N_0}} \\ &= \overbrace{\sum_{m=0}^{N_0/2-1} x[2m]\, e^{-j\frac{2\pi km}{N_0/2}}}^{\overline{P}_k} + e^{-j\frac{2\pi k}{N_0}} \cdot \overbrace{\sum_{m=0}^{N_0/2-1} x[2m+1]\, e^{-j\frac{2\pi km}{N_0/2}}}^{\overline{D}_k}, \quad k = 0,\ldots,N_0-1 \end{aligned} \tag{5.6.5}$$

La prima sommatoria rappresenta la trasformata discreta di una sequenza costituita dagli $N_0/2$ *campioni di indice pari di* $x[n]$, mentre la seconda sommatoria è la trasformata discreta degli $N_0/2$ *campioni di indice dispari*. Possiamo dire che questa scomposizione è "ricorsiva nell'ordine", nel senso che la trasformata di ordine N_0 è espressa come combinazione lineare di due trasformate di ordine $N_0/2$.

---

[2] Un *algoritmo* è una successione di passi od operazioni di elaborazione univocamente definite che, a partire da una serie di *dati* d'ingresso, nel nostro caso i valori della sequenza $x[n]$, fornisce i *dati di uscita*, qui i valori della trasformata discreta $\overline{X}_k$.

[3] L'acronimo FFT indica in realtà una *classe* di algoritmi efficienti per il calcolo della trasformata discreta di una sequenza periodica. Di seguito viene descritto solo uno di questi algoritmi che è noto come *algoritmo a decimazione nel tempo*.

Il numero di operazioni $N_{FFT}(N_0)$ necessario a calcolare la trasformata di ordine N_0 secondo questo nuovo criterio può allora essere espresso in maniera ugualmente *ricorsiva* sulla base di questa scomposizione:

$$\begin{aligned} N_{FFT}(N_0) &= N_{FFT}\left(\frac{N_0}{2}\right) + N_{FFT}\left(\frac{N_0}{2}\right) + 6N_0 + 2N_0 \\ &= 2\left[N_{FFT}\left(\frac{N_0}{2}\right) + 4N_0\right] \end{aligned} \tag{5.6.6}$$

avendo tenuto conto del fatto che, per ogni valore di k, è necessario moltiplicare $\overline{D}_k$ per un esponenziale complesso (precalcolato, 6 operazioni reali) e quindi effettuare la somma con $\overline{P}_k$ (2 operazioni reali). Questo procedimento di scomposizione può essere poi *ripetuto in modo ricorsivo*. Infatti $\overline{P}_k$ e $\overline{D}_k$ possono a loro volta essere scomposti suddividendo le sequenze rispettivamente $x[2m]$ e $x[2m+1]$ in due sottosequenze ciascuna di lunghezza $N_0/4$. Il calcolo di una trasformata di ordine $N_0/2$ comporta allora una complessità (basta sostituire $N_0/2$ a N_0 nella (5.6.6))

$$N_{FFT}\left(\frac{N_0}{2}\right) = 2\left[N_{FFT}\left(\frac{N_0}{4}\right) + \frac{4N_0}{2}\right] \tag{5.6.7}$$

Sostituendo quest'ultima nella relazione di partenza (5.6.6) si ha

$$N_{FFT}(N_0) = 2\left\{2\left[N_{FFT}\left(\frac{N_0}{4}\right) + \frac{4N_0}{2}\right] + 4N_0\right\} = 4\,N_{FFT}\left(\frac{N_0}{4}\right) + 2 \cdot 8N_0 \tag{5.6.8}$$

e, continuando a iterare dividendo per 2 progressivamente l'ordine di trasformazione, si ottiene:

$$\begin{aligned} N_{FFT}(N_0) &= 8\,N_{FFT}\left(\frac{N_0}{8}\right) + 3 \cdot 8N_0 \\ N_{FFT}(N_0) &= 16\,N_{FFT}\left(\frac{N_0}{16}\right) + 4 \cdot 8N_0 \\ &\vdots \\ N_{FFT}(N_0) &= N_0 \cdot N_{FFT}\left(\frac{N_0}{N_0}\right) + \log_2 N_0 \cdot 8N_0 \end{aligned} \tag{5.6.9}$$

Nell'ultima iterazione compare la quantità $N_{FFT}(1)$ che comporta banalmente una sola moltiplicazione ($\overline{X}_0 = x[0] \cdot e^{j0}$), per cui si ricava:

$$N_{FFT}(N_0) = 6N_0 + 8N_0 \log_2 N_0 \cong 8N_0 \log_2 N_0 \tag{5.6.10}$$

Questa relazione estremamente importante indica una complessità per l'algoritmo di FFT *notevolmente inferiore* alla complessità (quadratica) dell'algoritmo di trasformata discreta secondo la definizione. Il rapporto tra il numero di operazioni necessarie nei due casi è pari a

$$\frac{N_{TDF}(N_0)}{N_{FFT}(N_0)} = \frac{8N_0^2}{8N_0 \log_2 N_0} = \frac{N_0}{\log_2 N_0} \tag{5.6.11}$$

Con riferimento all'esempio precedente, se si usa il medesimo elaboratore per calcolare una FFT di ordine 1024, si otterrà un tempo di calcolo inferiore a quello prima calcolato del fattore

$$N_0 / \log_2 N_0 = 1024/10 \cong 100 \tag{5.6.12}$$

e di conseguenza la frequenza di campionamento massima per operare in tempo reale sarà

$$f_{c,FFT} = \frac{N_0}{\log_2 N_0} f_c \cong 100 f_c = 1.28 \text{ MHz} \tag{5.6.13}$$

che è molto maggiore di quella indicata nella (5.6.4). Dalla (5.6.11) è anche chiaro che il vantaggio che si consegue utilizzando l'algoritmo di calcolo FFT invece della trasformata secondo la definizione *aumenta al crescere di* N_0. Come già detto, gli stessi algoritmi utilizzati per il calcolo veloce della trasformata discreta possono essere anche utilizzati per il calcolo della *antitrasformata*, poiché le relazioni di trasformazione e di antitrasformazione (5.5.2-5.5.3) sono formalmente identiche, a parte l'inessenziale costante moltiplicativa $1/N_0$, e l'altrettanto ininfluente segno dell'argomento degli esponenziali.

L'algoritmo di FFT fu pubblicato nel 1965 da Cooley e Tukey; a questa data si fa risalire la nascita della moderna *elaborazione numerica dei segnali*. Attraverso la FFT, infatti, possono essere effettuate in maniera efficiente alcune operazioni fondamentali di analisi ed elaborazione dei segnali (analisi spettrale e filtraggio, come mostreremo brevemente nel seguito). Tali operazioni non furono realizzabili in pratica fino al momento dell'introduzione dell'algoritmo veloce, vista la ridotta velocità dei componenti elettronici e quindi dei calcolatori dell'epoca. Questo giustifica l'importanza centrale attribuita alla FFT nello sviluppo delle tecniche di elaborazione numerica.

5.6.2 Applicazioni dell'algoritmo di FFT: analisi spettrale

Una delle esigenze più frequenti nell'elaborazione dei segnali è quella di calcolare lo spettro di una sequenza data. Ovviamente, non si ha a disposizione un'*espressione analitica* dei valori della sequenza $x[n]$, ma solo i valori medesimi (ad esempio acquisiti mediante un convertitore A/D) in un intervallo finito, diciamo $0 \le n \le N-1$. Ciò che si desidererebbe calcolare è la trasformata di Fourier della sequenza aperiodica

$$\overline{X}(f) = \sum_{n=-\infty}^{+\infty} x[n]\, e^{-j2\pi nfT} \tag{5.6.14}$$

Ovviamente, questa trasformata non potrà essere calcolata per gli *infiniti* valori della variabile f in un periodo, cioè, ad esempio, nell'intervallo $[0, 1/T]$. Ci si accontenterà di ottenere il valore della trasformata per un numero finito di punti normalmente *equispaziati* nell'intervallo $[0, 1/T]$.

Quest'operazione può essere svolta in maniera *efficiente* tenendo conto della relazione di campionamento in frequenza (5.5.28) conseguente ad una periodicizzazione:

$$\overline{Y}_k = \frac{1}{N_0}\overline{X}\left(\frac{k}{N_0 T}\right) \quad , \quad k = 0, 1, \ldots, N-1 \tag{5.6.15}$$

e osservando che abbiamo a disposizione un algoritmo *veloce* per il calcolo delle trasformate discrete. Immaginiamo dunque di *periodicizzare* la sequenza data con periodo $N_0 = N$:

$$y[n] = \sum_{m=-\infty}^{+\infty} x[n - mN] \tag{5.6.16}$$

Poiché $y[n]$ è una sequenza periodica, possiamo calcolarne la trasformata discreta di Fourier $\overline{Y}_k$, $k = 0, 1, \ldots, N-1$ *mediante un algoritmo veloce*. Sfruttando la (5.6.15) possiamo poi ricavare i seguenti valori di $\overline{X}(f)$:

$$\overline{X}\left(\frac{k}{NT}\right) = N\overline{Y}_k \quad , \quad k = 0, 1, \ldots, N-1 \tag{5.6.17}$$

Siamo cioè riusciti a calcolare la funzione $\overline{X}(f)$ in N punti, e precisamente per le N frequenze *equispaziate nell'intervallo* $[0, 1/T]$:

$$f_k = \frac{k}{NT} \quad , \quad k = 0, 1, \ldots, N-1 \tag{5.6.18}$$

Esempio 5.14

Consideriamo la sequenza aperiodica $x[n]$ di Figura 5.36: essa ha durata finita e pari a $N = 3$. Consideriamo poi il segnale *periodico* $y[n]$ ottenuto periodicizzando con periodo 3 la sequenza $x[n]$ in accordo alla relazione (5.6.16), e rappresentiamo la sequenza $y[n]$ così ottenuta (Figura 5.37).

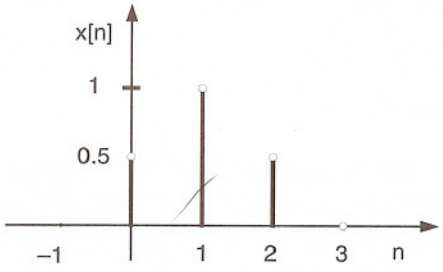


Figura 5.36 Sequenza aperiodica $x[n]$ a durata finita

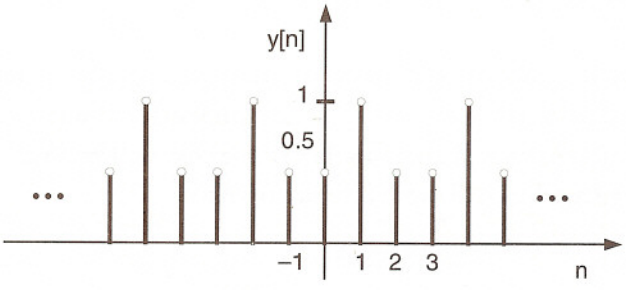


Figura 5.37 Sequenza periodica ottenuta dalla ripetizione di $x[n]$ per $N_0 = N = 3$

La trasformata di Fourier $\overline{X}(f)$ della sequenza originaria è

$$\overline{X}(f) = x[0] + x[1]\,e^{-j2\pi fT} + x[2]\,e^{-j2\pi 2fT} = e^{-j2\pi fT}\left[1 + \cos(2\pi fT)\right] \qquad \text{(E5.14.1)}$$

corrispondente allo spettro di ampiezza di Figura 5.38. Utilizzando il metodo appena descritto, si usa una trasformata discreta di ordine 3 e si ricavano i 3 "campioni" della funzione $\overline{X}(f)$ per le frequenze $f_k = 0, 1/3T, 2/3T$ rappresentati ancora in Figura 5.38. ❑

Spesso, la scelta $N_0 = N$ porta a una conoscenza insufficiente della funzione $\overline{X}(f)$, nel senso che gli N campioni della trasformata discreta sono troppo pochi per ottenere una stima sufficientemente accurata dell'andamento della

$\overline{X}(f)$ stessa, come nel caso della Figura 5.38. Per evitare questo inconveniente è possibile usare per il parametro N_0 un valore *maggiore* della durata della sequenza N. Il segnale $y[n]$ viene ottenuto cioè dalla ripetizione di una sequenza finita di durata N_0 formata dagli N campioni della $x[n]$ a cui vengono posposti $N_0 - N$ campioni *nulli*. Si può calcolare adesso la FFT di ordine $N_0 > N$ della sequenza periodica che nel periodo base vale

$$y[n] = \begin{cases} x[n] & 0 \le n \le N-1 \\ 0 & N \le n \le N_0 - 1 \end{cases} \tag{5.6.19}$$

cioè di una sequenza riempita con campioni nulli. Questa operazione prende il nome di *zero-padding* o *riempimento con zeri*. Una volta calcolata la trasformata discreta, vengono ricavati quindi $N_0 > N$ valori della $\overline{X}(f)$ secondo la consueta relazione di campionamento in frequenza:

$$\overline{X}\left(\frac{k}{N_0 T}\right) = N_0 \overline{Y}_k \quad , \quad k = 0, 1, \ldots, N_0 - 1 \tag{5.6.20}$$

Se si fa crescere il valore del parametro N_0, ossia si aggiungono molti campioni nulli aumentando l'ordine della FFT, si può aumentare il numero di *campioni* della funzione $\overline{X}(f)$, cioè si aumenta la *risoluzione* dell'analisi spettrale della sequenza $x[n]$ aperiodica e di durata finita originaria.

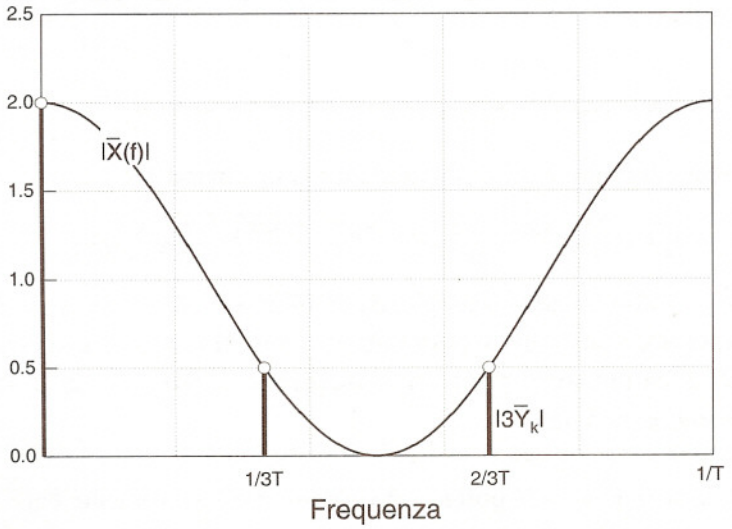


Figura 5.38 Trasformata discreta della sequenza $x[n]$ dell'Esempio 5.14

Esempio 5.15

Ripetiamo il calcolo dell'Esempio 5.11, stavolta però con *zero-padding* fino a $N_0 = 8$. La sequenza periodicizzata dopo il riempimento con 5 zeri è quella di Figura 5.39. Calcolando la FFT della sequenza dopo lo zero-padding ricaveremo allora 8 valori della trasformata $\overline{X}(f)$ secondo la relazione

$$\overline{X}\left(\frac{k}{8T}\right) = 8\,\overline{Y}_k \quad , \quad k = 0, 1, ..., 7 \qquad \text{(E5.15.1)}$$

Figura 5.39 Sequenza periodica ottenuta dalla ripetizione di $x[n]$ per $N_0 = 8$

Nella Figura 5.40 viene riproposto l'andamento del modulo di $\overline{X}(f)$ e dei suoi otto campioni calcolati con il procedimento di zero-padding ed FFT appena descritto. L'andamento dello spettro di ampiezza $|\overline{X}(f)|$ viene ora ricavato con una *risoluzione* maggiore che nel caso precedente. Vale la pena osservare che l'operazione di *periodicizzazione* concettualmente necessaria prima di calcolare la trasformata discreta *non viene eseguita in pratica.* Ciò che serve infatti per il calcolo della trasformata sono soltanto gli N_0 campioni della sequenza $y[n]$ nel periodo base $0 \le n \le N_0 - 1$. Viceversa, l'operazione di zero-padding è *fondamentale* per costruire correttamente $y[n]$ a partire da $x[n]$, e deve essere effettivamente eseguita prima di calcolare la trasformata discreta $\overline{Y}_k$. ❑

5.6.3 Applicazioni dell'algoritmo FFT: convoluzione veloce

Un'operazione di elaborazione dei segnali che si presenta molto frequentemente è quella del calcolo di una *somma di convoluzione* tra sequenze di durata finita (tipico problema di *filtraggio*, come sarà chiarito nel Capitolo 6). Siano dunque $x[n]$ e $y[n]$ due sequenze aperiodiche (per semplicità causali) a durata finita pari a N.

La convoluzione (lineare) tra le due sequenze è

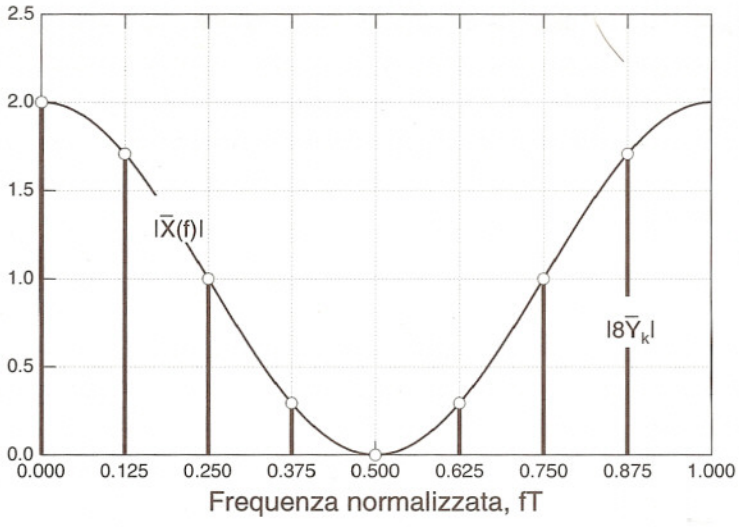


Figura 5.40 Risultato dell'operazione di zero-padding e FFT

$$z[n] = x[n] \otimes y[n] = \sum_{m=-\infty}^{+\infty} x[m]\, y[n-m] = \sum_{m=0}^{N-1} x[m]\, y[n-m] \tag{5.6.21}$$

Poiché $x[n]$ e $y[n]$ hanno durata N, cioè hanno solo N campioni non nulli, è immediato rendersi conto che $z[n]$ ha durata pari a $2N-1$ (si estende cioè da 0 a $2N-2$): un esempio è illustrato in Figura 5.41 per $N=4$.

Cerchiamo ora di valutare il grado di complessità del calcolo (5.6.21) della convoluzione eseguita attraverso la definizione. Nella Tabella 5.2 abbiamo riportato per esteso tutte le operazioni necessarie al calcolo di tutti i valori di $z[n]$, con $0 \le n \le 2N-2$, tenendo conto che le sequenze $x[n]$ e $y[n]$ sono entrambe nulle quando $n<0$ o $n \ge N$. Il numero totale di operazioni $N_{conv}(N)$ si può valutare notando che le operazioni necessarie al calcolo dei termini $z[0]$ e $z[2N-2]$ sono in ugual numero, e così per i termini $z[1]$ e $z[2N-3]$, $z[2]$ e $z[2N-4]$, ecc., fino ad arrivare al termine $z[N-1]$ che non ha nessun "partner". I primi due termini richiedono 1 sola operazione ciascuno, i secondi due 3 operazioni, i successivi 5 operazioni e così via, fino al termine $z[N-1]$ che richiede $2N-1$ operazioni:

$$N_{conv}(N) = 2 \cdot [1+3+5+\ldots+(2(N-2)+1)] + (2N-1) =$$
$$= 2 \cdot [1+3+5+\ldots+(2(N-2)+1)+(2(N-1)+1)] - (2N-1) =$$

$$= 2\sum_{n=0}^{N-1}(2n+1) - (2N-1) = 4\frac{N(N-1)}{2} + 2N - 2N + 1$$

$$= 2N^2 - 2N + 1 \cong 2N^2 \qquad (5.6.22)$$

cioè la complessità è di tipo *quadratico* nella durata delle sequenze.

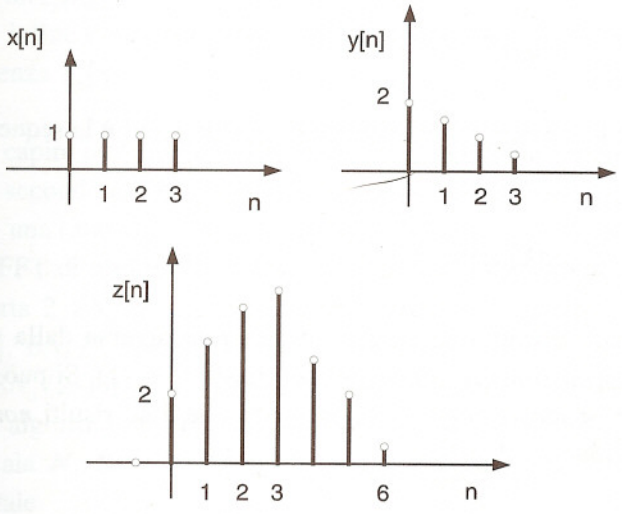


Figura 5.41 Convoluzione tra le sequenze $x[n]$ e $y[n]$

Tabella 5.2 Operazioni necessarie al calcolo di una convoluzione

$$z[0] = x[0] \cdot y[0]$$
$$z[1] = x[0] \cdot y[1] + x[1] \cdot y[0]$$
$$z[2] = x[0] \cdot y[2] + x[1] \cdot y[1] + x[2] \cdot y[0]$$
$$\vdots$$
$$z[N-2] = x[0] \cdot y[N-2] + x[1] \cdot y[N-3] + \ldots + x[N-2] \cdot y[0]$$
$$z[N-1] = x[0] \cdot y[N-1] + x[1] \cdot y[N-2] + x[2] \cdot y[N-3] + \ldots + x[N-1] \cdot y[0]$$
$$z[N] = x[1] \cdot y[N-1] + x[2] \cdot y[N-2] + \ldots + x[N-1] \cdot y[0]$$
$$\vdots$$
$$z[2N-4] = x[N-3] \cdot y[N-1] + x[N-2] \cdot y[N-2] + x[N-1] \cdot y[N-3]$$
$$z[2N-3] = x[N-2] \cdot y[N-1] + x[N-1] \cdot y[N-2]$$
$$z[2N-2] = x[N-1] \cdot y[N-1]$$

La somma di convoluzione tra $x[n]$ ed $y[n]$ può anche essere determinata per altra via. Supponiamo di periodicizzare le due sequenze con periodo N_0:

$$x_P[n] = \sum_{m=-\infty}^{+\infty} x[n - mN_0] \tag{5.6.23}$$

$$y_P[n] = \sum_{m=-\infty}^{+\infty} y[n - mN_0] \tag{5.6.24}$$

Sappiamo che la convoluzione *ciclica* tra $x_P[n]$ e $y_P[n]$ è la sequenza di periodo N_0

$$z_p[n] = \frac{1}{N_0} \sum_{k=0}^{N_0-1} x_P[k]\, y_P[n-k] \tag{5.6.25}$$

Questa sequenza (periodica) $z_p[n]$ è in generale diversa dalla sequenza $z[n]$ prodotta dall'operazione di convoluzione lineare (5.6.21). Si può però scegliere l'intervallo di periodicizzazione N_0 in modo che esso risulti *non minore* della durata della sequenza $z[n]$

$$N_0 \geq 2N - 1 \tag{5.6.26}$$

Allora è chiaro che l'operazione di convoluzione *ciclica* fra le due sequenze *periodicizzate* $x_P[n]$ e $y_P[n]$ dà una sequenza $z_p[n]$ che all'interno del "periodo base" $0 \leq n \leq N_0 - 1$ *coincide con il risultato della convoluzione aperiodica* a meno del fattore $1/N_0$:

$$z[n] = \begin{cases} N_0\, z_p[n] & n = 0, 1, \ldots, 2N-2 \\ 0 & \text{altrove} \end{cases} \tag{5.6.27}$$

D'altronde, dal teorema della convoluzione (5.5.19), sappiamo che la trasformata discreta di $z_p[n]$ è

$$\overline{Z}_{p_k} = \overline{X}_{p_k} \overline{Y}_{p_k} \tag{5.6.28}$$

e questo suggerisce un metodo indiretto per il calcolo della convoluzione in ambito frequenziale, costituito dai seguenti passi:

- si sceglie per il parametro N_0 il minimo valore che soddisfa la disuguaglianza (5.6.26), cioè $N_0 = 2N - 1$;

- si calcolano mediante FFT le trasformate discrete $\overline{X}_{p_k}$ e $\overline{Y}_{p_k}$ di ordine N_0 delle sequenze periodicizzate $x_P[n]$ e $y_P[n]$;
- si calcola la trasformata $\overline{Z}_{p_k}$ della convoluzione ciclica $z_p[n]$ delle sequenze $x_P[n]$ e $y_P[n]$ come prodotto di $\overline{X}_{p_k}$ e $\overline{Y}_{p_k}$;
- si calcola mediante FFT l'antitrasformata di $\overline{Z}_{p_k}$ ricavando così i campioni di un periodo di $z_p[n]$;
- si ricavano infine i campioni non nulli della sequenza $z[n]$ moltiplicando per N_0 la sequenza $z_p[n]$.

Cerchiamo di capire quando questo procedimento è più conveniente rispetto alla convoluzione secondo definizione. Si può dimostrare che per sequenze *reali* si può effettuare una (anti-)FFT di una sequenza di lunghezza N_0 usando opportunamente una FFT di ordine $N_0/2$. Quindi, secondo la (5.6.10), la convoluzione veloce comporta $2 \cdot 8N_0/2 \cdot \log_2 N_0/2$ operazioni per il calcolo delle due FFT di $x_P[n]$ e $y_P[n]$, $6N_0/2$ operazioni ($N_0/2$ moltiplicazioni complesse) per il calcolo di $\overline{Z}_{p_k}$, $8N_0/2 \cdot \log_2 N_0/2$ operazioni per calcolare l'antitrasformata $z_p[n]$ di $\overline{Z}_{p_k}$ con un algoritmo di FFT inversa, ed infine N_0 operazioni per recuperare il fattore di scala N_0. Se $N_0 = 2N-1$, e trascurando termini *lineari* in N_0, sono richieste in totale

$$N_{conv,FFT}(N) \cong 12N_0/2\log_2(N_0/2) \cong 12N\log_2(N) \qquad (5.6.29)$$

operazioni ($N >> 1$). Questo risultato deve essere confrontato con la (5.6.22): è sufficiente che sia $N \geq 32$ affinché il metodo che fa uso di FFT diventi più efficiente del calcolo diretto della convoluzione mediante la definizione.

Esempio 5.16

Riprendiamo in considerazione le sequenze della Figura 5.41 e cerchiamo di applicare il metodo della convoluzione veloce *senza tener conto della condizione* (5.6.26). Fissiamo allora, *in modo errato* $N_0 = 5$ (sarebbe richiesto $N_0 \geq 7$). Le sequenze periodicizzate in questo modo sono rappresentate in Figura 5.42a. La relativa convoluzione *ciclica*, scalata del fattore $N_0 = 5$, è poi confrontata in Figura 5.42b con la convoluzione *lineare* che si desidererebbe ottenere. È chiaro che non può sussistere l'uguaglianza, segnatamente perché la convoluzione lineare ha durata pari a 7, mentre la convoluzione ciclica è periodica di periodo 5 e non può quindi fornire i 7 campioni distinti della convoluzione lineare.

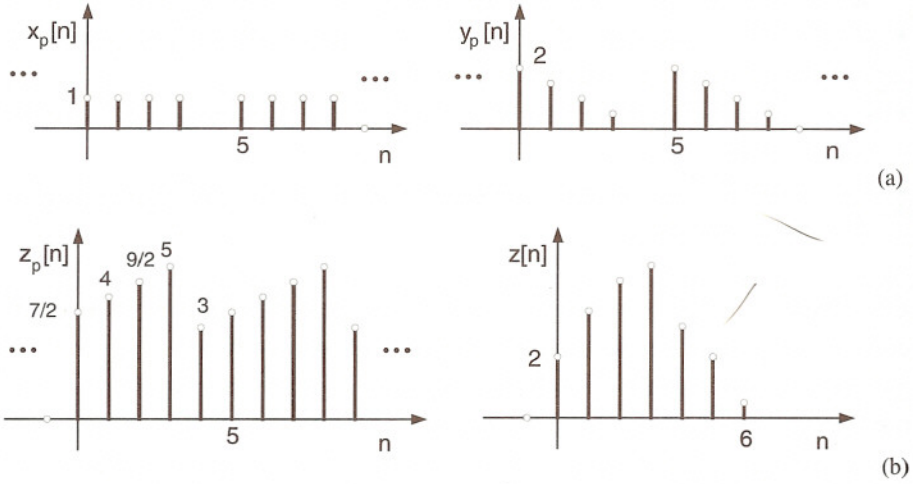


Figura 5.42 Sequenze di Figura 5.41 periodicizzate (a), convoluzione ciclica e lineare (b)

❑

5.7 Riassunto delle caratteristiche delle trasformate di Fourier

La Figura 5.43 riassume le caratteristiche delle descrizioni frequenziali (spettri) dei segnali a tempo continuo e a tempo discreto, periodici e aperiodici.

Ogniqualvolta il segnale è *periodico* nel tempo, esso possiede uno spettro *discreto*. Viceversa, se il segnale è *discreto* nel tempo, esso possiede un spettro *periodico*. Questo è l'ennesimo riflesso della *dualità* dei domini di tempo e frequenza. Da quest'ultimo punto di vista, è interessante notare che il segnale discreto aperiodico $x[n]$ è in pratica *la successione dei coefficienti di Fourier* dell'espansione in serie della funzione $\overline{X}(f)$, periodica nella variabile continua frequenza; ciò in piena dualità rispetto al caso del segnale periodico nel tempo continuo $x(t)$ con la sua propria successione discreta dei coefficienti di Fourier X_k:

$$\begin{aligned} x[n] &= \frac{1}{1/T}\int_{-1/2T}^{1/2T} \overline{X}(f)\, e^{j2\pi nfT} df \\ \overline{X}(f) &= \sum_{n=-\infty}^{\infty} x[n] e^{-j2\pi nfT} \end{aligned} \qquad \begin{aligned} X_k &= \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t) e^{-j2\pi k f_0 t} dt \\ x(t) &= \sum_{k=-\infty}^{\infty} X_k e^{j2\pi k f_0 t} \end{aligned} \qquad (5.7.1)$$

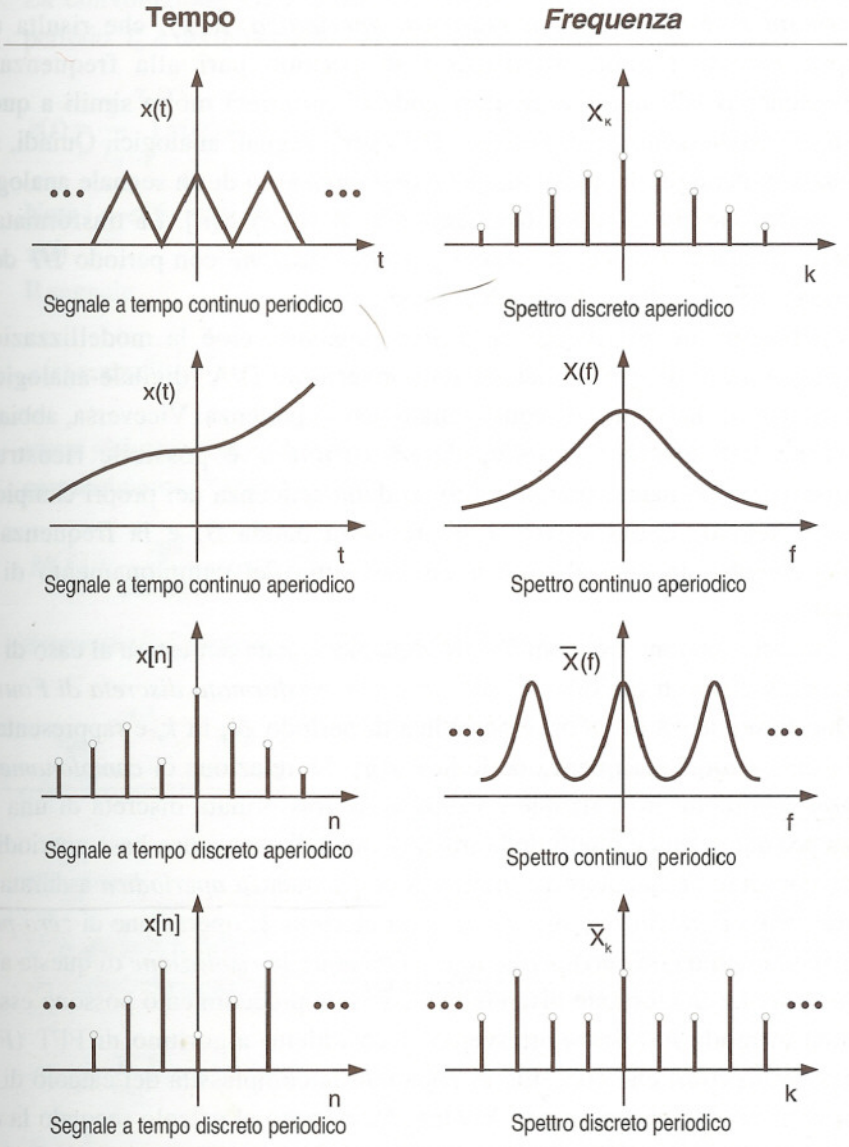


Figura 5.43 Tavola sinottica delle caratteristiche di segnali e spettri

Sommario

In questo capitolo sono stati ripresi in considerazione ed estesi ai segnali a *tempo discreto* alcuni concetti relativi all'analisi di Fourier già esaminati nei precedenti

capitoli per i segnali a tempo continuo. Per prima cosa è stata definita la *trasformata di Fourier di una sequenza aperiodica* $\overline{X}(f)$, che risulta una funzione *periodica* nella frequenza f di periodo pari alla frequenza di campionamento $1/T$, ma che peraltro gode di proprietà molto simili a quelle della trasformata continua di Fourier $X(f)$ per i segnali analogici. Quindi, si è esaminata in dettaglio la questione del *campionamento* di un segnale analogico $x(t)$, operazione che produce una sequenza di valori $x[n]$. La trasformata di questa sequenza si ottiene attraverso *periodicizzazione* con periodo $1/T$ della trasformata del segnale analogico di partenza.

L'operazione di *interpolazione a mantenimento*, cioè la modellizzazione dell'operazione svolta in pratica da un convertitore D/A (digitale-analogico), non consente di ricostruire il segnale analogico di partenza. Viceversa, abbiamo dimostrato che usando un *interpolatore cardinale* è possibile ricostruire esattamente un segnale a tempo continuo dalla sequenza dei propri campioni, purché il segnale abbia spettro limitato nella banda B, e la frequenza di campionamento sia pari almeno a $2B$ (teorema del campionamento di C. Shannon).

La rappresentazione frequenziale di sequenze è stata poi estesa al caso di sequenze *periodiche* di periodo N_0 definendo la *trasformata discreta di Fourier* $\overline{X}_k$. Questa sequenza di valori è periodica di periodo N_0 in k, e rappresenta lo spettro *discreto* della sequenza periodica $x[n]$. La relazione di *campionamento in frequenza* mette in relazione i valori della trasformata discreta di una sequenza periodicizzata e quelli della trasformata della sequenza-base aperiodica. Questo consente di calcolare lo spettro di una sequenza *aperiodica* a durata finita attraverso il calcolo di una trasformata discreta. L'operazione di *zero-padding* (riempimento con zeri) permette di aumentare la *risoluzione* di questa analisi spettrale. Le trasformate discrete insite in tale procedimento possono essere calcolate in modo *efficiente* attraverso il cosiddetto algoritmo di FFT (*Fast Fourier Transform*) che consente di abbattere la complessità del calcolo di un trasformata discreta di un fattore $N_0 / \log_2 N_0$ rispetto al calcolo secondo la definizione. Sfruttando la FFT è anche possibile calcolare *somme di convoluzione* tra sequenze a durata finita in maniera veloce.

Esercizi proposti

5.1 Ricavare la trasformata di Fourier della sequenza

$$x[n] = a^{|n|} \quad , \quad 0 \le a < 1$$

5.2 La convoluzione ciclica tra due segnali a tempo *continuo* periodici di periodo T_0 è definita come

$$z(t) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(\alpha) y(t-\alpha)\, d\alpha$$

Noti i coefficienti di Fourier di $x(t)$ e $y(t)$, ricavare i coefficienti Z_k di $z(t)$.

5.3 Il segnale

$$x(t) = \text{sinc}^2\left(\frac{t}{T}\right)$$

viene elaborato secondo lo schema di Figura 5.44, in cui l'interpolatore è a mantenimento. Sapendo che

$$H(f) = \frac{\pi f T}{\text{sen}(\pi f T)} \text{rect}(fT)$$

determinare l'espressione del segnale di uscita $y(t)$.

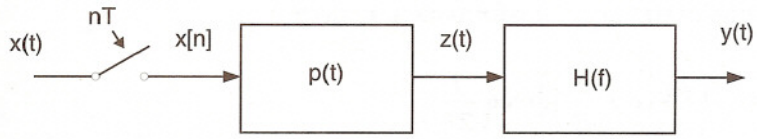


Figura 5.44

5.4 Una sequenza periodica a tempo discreto può essere espressa attraverso la serie discreta (antitrasformata) di Fourier

$$x[n] = \sum_{k=0}^{N_0-1} \overline{X}_k e^{j2\pi kn/N_0}$$

Trovare l'espressione della trasformata *aperiodica* $\overline{X}(f)$ della sequenza $x[n]$ per $f \in [-1/2T, 1/2T]$.

5.5 Applicando il teorema della funzione somma, calcolare la trasformata della sequenza

$$x[n] = \begin{cases} 1 - \dfrac{|n|}{N} & -N \le n \le N \\ 0 & \text{altrove} \end{cases}$$

5.6 Disegnare i segnali ricostruiti con interpolazione a mantenimento e interpolazione lineare rispettivamente nellea Figure 5.21 e nella Figura 5.29. Ripetere per il segnale di Figura 5.30 (l'espressione della trasformata in questa figura è del tipo $X(f) = (T/2)[1 + \cos(2\pi fT)]\text{rect}(fT)$).

5.7 Un segnale $x(t)$ periodico di periodo T_0 viene campionato con periodo $T = T_0 / N_0$ ottenendo la sequenza periodica $x[n]$. Determinare la relazione tra i coefficienti di Fourier X_k di $x(t)$ e la trasformata discreta di Fourier $\overline{X}_k$ di $x[n]$

5.8 Dire se il sistema monodimensionale *a tempo continuo* costituito dalla cascata di campionatore ideale e interpolatore a mantenimento di Figura 5.18 è lineare e/o stazionario.

5.9 Da un segnale $x(t)$ periodico di periodo T_0 viene ricavato il segnale $y(t)$ *aperiodico* con una operazione di *troncamento* in un periodo:

$$y(t) = x(t)\text{rect}(t / T_0)$$

Esprimere la trasformata di Fourier $Y(f)$ di $y(t)$ mediante i coefficienti di Fourier X_k di $x(t)$, e spiegare perché questo esercizio è pertinente a questo capitolo.

5.10 Il segnale $x(t)$ la cui trasformata di Fourier è

$$X(f) = \left(1 + \cos\left(\frac{\pi f}{B}\right)\right)\text{rect}\left(\frac{f}{2B}\right)$$

viene campionato con frequenza $f_c = B$, e quindi interpolato con un interpolatore cardinale. Trovare l'espressione del segnale $\hat{x}(t)$ così ottenuto e tracciarne un grafico.

5.11 Calcolare la trasformata di Fourier $\overline{X}(f)$ della sequenza

$$x[n] = \begin{cases} 16 \cdot \left(\frac{1}{2}\right)^n & n \geq 4 \\ 1 & |n| \leq 3 \\ 16 \cdot 2^n & n \leq -4 \end{cases}$$

5.12 Determinare e rappresentare lo spettro della sequenza $x[n]$ ottenuta campionando con frequenza $1/T$ il segnale $x(t) = \text{sinc}(Bt)\sin(\pi Bt)$ con $B \leq 1/T \leq 2B$. È possibile ricostruire il segnale originario $x(t)$ a partire dai campioni della sequenza $x[n]$?

5.13 Il segnale $x(t) = e^{-|t|/\alpha}$ viene campionato con frequenza di campionamento $f_c = 1/T$. Disegnare gli spettri di ampiezza e fase della sequenza $x[n]$ così ottenuta.

5.14 Dimostrare l'uguaglianza

$$\sum_{k=-\infty}^{\infty} \frac{\alpha/T}{1 + j2\pi\left(f - \frac{k}{T}\right)\alpha} = \frac{1}{1 - e^{-\left(\frac{1}{\alpha} + j2\pi f\right)T}}$$

5.15 Calcolare e rappresentare la somma di convoluzione $z[n]$ tra le sequenze $x[n] = u[n+1] - u[n-2]$ e $y[n] = u[n+2] - u[n-3]$.

$$x(t) = \sum_k X_k e^{+2\pi k t/T_0}$$

$$y(t) = \sum_k X_k e^{+2\pi k t/T_0} \cdot \mathrm{rect}\left(\frac{t}{T_0}\right)$$

$$Y(f) = \sum_k X_{-k} \cdot T_0 \,\mathrm{sinc}\left(T_0\left(f - \frac{k}{T_0}\right)\right) =$$

$$= \sum_k X_k \frac{1}{f_0} \,\mathrm{sinc}\left(\frac{(f - k f_0)}{f_0}\right)$$

5.2 Rappresentazione dei segnali aperiodici a tempo discreto nel dominio della frequenza

5.2.1 Trasformata di Fourier di una sequenza

Dai capitoli precedenti risulta chiara l'utilità della rappresentazione dei segnali analogici nel dominio della frequenza, sia come strumento di analisi e di sintesi, sia in congiunzione con lo studio dei sistemi lineari stazionari. Vogliamo ora estendere la rappresentazione frequenziale anche ai segnali a tempo discreto, e per far questo seguiremo un cammino inverso rispetto a quello percorso per i segnali a tempo continuo: partiremo cioè dalle sequenze *aperiodiche*. Gli aspetti di carattere concettuale insiti alla rappresentazione dei segnali nel dominio frequenziale restano i medesimi, e sono ormai patrimonio acquisito del lettore.

La rappresentazione di una sequenza aperiodica $x[n]$ in campo frequenziale analoga alla trasformata continua di Fourier di un segnale analogico aperiodico (3.1.9) è la *trasformata di Fourier della sequenza* $x[n]$ definita da

$$\overline{X}(F) \triangleq \sum_{n=-\infty}^{\infty} x[n]\, e^{-j2\pi nF} \tag{5.2.1}$$

Questa trasformata è evidentemente una funzione complessa della variabile F che è la cosiddetta *frequenza normalizzata* (si confronti con la (5.1.10)). La stessa trasformata viene talvolta espressa in funzione di una *pulsazione* normalizzata $\Omega = 2\pi F$:

$$\overline{X}(\Omega) \triangleq \sum_{n=-\infty}^{\infty} x[n]\, e^{-jn\Omega} \tag{5.2.2}$$

Come vedremo più avanti, la $\overline{X}(F)$ definita dall'equazione di *analisi* (5.2.1) continua ad avere il significato di *spettro* del segnale dato, ma si differenzia dalla familiare trasformata continua $X(f)$ per segnali analogici in alcune proprietà. Osserviamo preliminarmente che la trasformata di una sequenza è una funzione *periodica in F di periodo 1*:

$$\overline{X}(F+1) = \sum_{n=-\infty}^{\infty} x[n]\, e^{-j2\pi n(F+1)} = \sum_{n=-\infty}^{\infty} x[n]\, e^{-j2\pi nF} e^{-j2\pi n}$$

$$= \sum_{n=-\infty}^{\infty} x[n]\, e^{-j2\pi nF} = \overline{X}(F) \tag{5.2.3}$$

Quindi $\overline{X}(F)$ è completamente nota se è noto il suo andamento in un intervallo

6

Sistemi monodimensionali a tempo discreto

6.1 Caratterizzazione dei sistemi a tempo discreto

La maggior parte delle considerazioni che sono già state fatte nel Paragrafo 4.1 a proposito dei sistemi a tempo continuo possono essere ripetute per i sistemi a *tempo discreto.* Un sistema monodimensionale a tempo discreto è un dispositivo, un apparato, un programma per calcolatore che elabora una sequenza *d'ingresso* $x[n]$ e genera una sequenza *di uscita* $y[n]$ (Figura 6.1). L'unica differenza notevole dal punto di vista concettuale che abbiamo introdotto rispetto al caso del Paragrafo 4.1 è la menzione esplicita di un *programma per calcolatore.* Infatti, secondo la discussione del Paragrafo 5.1, i segnali a tempo discreto possono essere elaborati da circuiti elettronici digitali programmabili (microprocessori) il cui funzionamento è regolato da un *programma.* Si può allora identificare il sistema che trasforma la sequenza d'ingresso in quella di uscita con il solo *programma* del circuito, piuttosto che con l'insieme di dispositivi e programma (hardware e software) che permette effettivamente di realizzare l'elaborazione stessa.

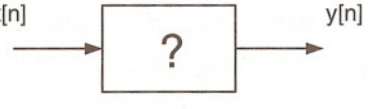


Figura 6.1 Sistema monodimensionale a tempo discreto

Stanti le forti analogie tra lo studio dei sistemi a tempo continuo già svolto nel Capitolo 4 e quello dei sistemi a tempo discreto, in questo paragrafo passeremo

brevemente in rassegna quelle definizioni e proprietà che sono già state discusse in maniera dettagliata per i sistemi a tempo continuo, limitandoci a sottolineare le (molte) similitudini e le (poche) diversità. Per maggiori dettagli e commenti su queste proprietà, rimandiamo il lettore al Paragrafo 4.1 relativo ai sistemi a tempo continuo.

6.1.1 Proprietà dei sistemi monodimensionali a tempo discreto

Identifichiamo dunque un sistema a tempo discreto con la trasformazione che viene eseguita sull'*eccitazione* $x[n]$ per fornire la *risposta* $y[n]$:

$$y[n] = \mathcal{T}[x[m]; n] \tag{6.1.1a}$$

o anche, quando non si corre alcun rischio di ambiguità,

$$y[n] = \mathcal{T}[x[n]] \tag{6.1.1b}$$

Un generico sistema monodimensionale a tempo discreto è:

- *lineare* se a esso è applicabile il principio di sovrapposizione degli effetti, cioè se

$$\mathcal{T}[\alpha x_1[n] + \beta x_2[n]] = \alpha \mathcal{T}[x_1[n]] + \beta \mathcal{T}[x_2[n]] \tag{6.1.2}$$

comunque siano fissate le sequenze $x_1[n]$ e $x_2[n]$ e i coefficienti α e β della loro combinazione lineare;

- *stazionario* o *invariante nel tempo* se una traslazione della sequenza in ingresso comporta una traslazione della stessa entità della sequenza di uscita, cioè

$$\mathcal{T}[x[n - n_0]] = y[n - n_0] \tag{6.1.3}$$

per ogni valore del parametro *intero* n_0;

- *causale* se la sequenza di uscita all'istante generico n non dipende dai valori assunti dalla sequenza di ingresso a istanti successivi a n:

$$y[n] = \mathcal{T}[x[m], m \leq n; n] = \mathcal{T}[x[m]u[n - m]; n] \tag{6.1.4}$$

- *stabile* secondo il criterio BIBO (*Bounded-Input, Bounded-Output*) se, per qualunque sequenza d'ingresso a valori limitati, tale cioè che

$$|x[n]| \leq K \quad \forall n \tag{6.1.5}$$

si ottiene una sequenza di uscita a sua volta a valori limitati, cioè

$$|x[n]| \leq K \Rightarrow |y[n]| \leq M \quad \forall n \tag{6.1.6}$$

• *istantaneo* o *senza memoria* se la sequenza di uscita all'istante generico n dipende solo dal valore assunto dalla sequenza di ingresso al medesimo istante:

$$y[n] = \mathcal{T}[x[m], m = n; n] = \mathcal{T}[x[m]\delta[n-m]; n] \tag{6.1.7}$$

• *invertibile* se è possibile trovare un secondo sistema $\mathcal{T}^{-1}[\cdot]$ tale che, per qualunque segnale di ingresso $x[n]$, si abbia

$$x[n] = \mathcal{T}^{-1}[y[n]] \tag{6.1.8}$$

6.1.2 Sistemi lineari e stazionari a tempo discreto

Rivolgiamo adesso la nostra attenzione alla classe dei sistemi lineari e stazionari (SLS) a tempo discreto. Come già dimostrato nel Paragrafo 4.1 per i sistemi analogici, un SLS a tempo discreto è caratterizzato completamente dalla conoscenza della *risposta impulsiva* $h[n]$ definita dalla relazione

$$h[n] = \mathcal{T}[\delta[n]] \tag{6.1.9}$$

Infatti, nota $h[n]$, la sequenza di uscita $y[n]$ del sistema avente in ingresso la sequenza $x[n]$ è

$$y[n] = \mathcal{T}[x[n]] = \mathcal{T}\left[\sum_{k=-\infty}^{+\infty} x[k]\,\delta[n-k]\right] \tag{6.1.10}$$

dalla quale, per la proprietà di *linearità*, si ricava

$$y[n] = \sum_{k=-\infty}^{+\infty} x[k]\,\mathcal{T}[\delta[n-k]] \tag{6.1.11}$$

Per la *stazionarietà* del sistema, poi, $\mathcal{T}[\delta[n-k]] = h[n-k]$ e la relazione precedente diviene

$$y[n] = \sum_{k=-\infty}^{+\infty} x[k]\,h[n-k] = x[n] \otimes h[n] \tag{6.1.12}$$

e cioè la sequenza di uscita di un SLS è la *somma di convoluzione* fra la sequenza di ingresso e la risposta impulsiva del sistema stesso.

Sulla falsariga di quanto visto per i sistemi analogici nel Paragrafo 4.2, è possibile dimostrare che condizione necessaria e sufficiente per la *stabilità in senso BIBO* di un SLS è la *assoluta sommabilità* della sua risposta impulsiva:

$$\sum_{k=-\infty}^{+\infty} |h[k]| < +\infty \tag{6.1.13}$$

e che inoltre un SLS è *causale* se e solo se la sua risposta impulsiva è una *sequenza causale*:

$$h[n] = h[n]u[n] \tag{6.1.14}$$

cioè $h[n] = 0$ se $n < 0$.

Esempio 6.1

Consideriamo il sistema, detto *filtro a media mobile*, che effettua la seguente trasformazione sulla sequenza di ingresso $x[n]$:

$$y[n] = \frac{1}{N} \sum_{k=n-N+1}^{n} x[k] \tag{E6.1.1}$$

con N dato. Si può verificare facilmente che tale sistema è *lineare* per la linearità dell'operazione di somma che lo definisce, ed è anche *stazionario*. Infatti

$$\mathcal{T}\{x[n-n_0]\} = \frac{1}{N} \sum_{k=n-N+1}^{n} x[k-n_0] = \frac{1}{N} \sum_{m=n-N+1-n_0}^{n-n_0} x[m] = y[n-n_0] \tag{E6.1.2}$$

Prima di procedere nell'analisi del sistema, cerchiamo di capire il tipo di funzione svolta dallo stesso (e di giustificarne quindi il nome). La Figura 6.2a riassume, per il caso particolare $N = 4$, le operazioni che si devono compiere sul segnale di ingresso per ottenere il valore $y[n^*]$ della sequenza di uscita al generico istante n^* (nella figura, $n^* = 4$). Secondo la relazione (E6.1.1), il valore $y[n^*]$ si ottiene come *media aritmetica* degli ultimi N valori della sequenza di ingresso $x[n]$, presi cioè a partire dall'istante n^* verso tempi decrescenti. Questi campioni sono i quattro contenuti all'interno della "finestra" ombreggiata in Figura 6.2a. Da questa descrizione e dalla figura si comprende anche che il sistema è *causale*; nella Figura 6.2a è pure messo in evidenza il meccanismo secondo il quale verrà calcolato il valore dell'uscita (indicato a linea tratteggiata) all'istante n^*+1: la "finestra" di N campioni della sequenza d'ingresso viene traslata di un passo in avanti (linea tratteggiata), e viene calcolata la media aritmetica dei nuovi N campioni contenuti nella finestra stessa (media *mobile*).

Poiché dunque il sistema è lineare e anche stazionario, è possibile calcolarne la risposta impulsiva che, per definizione, è data da

$$h[n] = \mathcal{T}[\delta[n]] = \frac{1}{N} \sum_{k=n-N+1}^{n} \delta[n] = \frac{1}{N} \left[\sum_{k=-\infty}^{n} \delta[k] - \sum_{k=-\infty}^{n-N} \delta[k] \right] \tag{E6.1.3}$$

Osservando che

$$u[n] = \sum_{k=-\infty}^{n} \delta[k] \tag{E6.1.4}$$

si può riscrivere la (E6.1.3) nella forma

$$h[n] = \frac{1}{N}(u[n] - u[n-N]) \tag{E6.1.5}$$

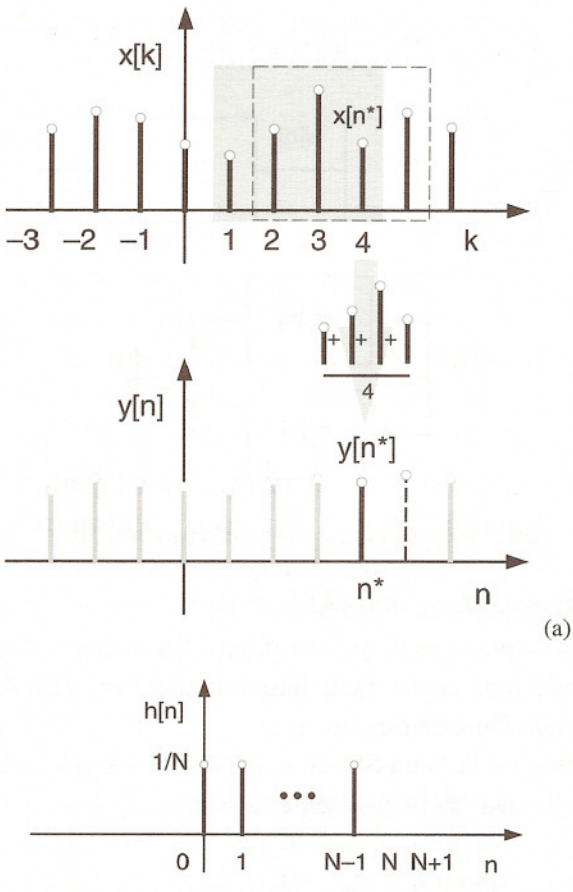


Figura 6.2 Funzionamento (a) e risposta impulsiva (b) del filtro a media mobile

da cui si vede che la risposta impulsiva del filtro a media mobile è l'impulso rettangolare *causale* di ampiezza $1/N$ raffigurato in Figura 6.2b. Il lettore può autonomamente ricalcolare l'andamento di $h[n]$ senza usare la relazione (E6.1.4), ma ricavando direttamente la risposta impulsiva attraverso il procedimento descritto nella Figura 6.2a. ❑

Quando due SLS aventi risposte impulsive $h_1[n]$ e $h_2[n]$ sono connessi in cascata o in parallelo, cioè rispettivamente come nelle Figure 6.3a e 6.3b, le relative risposte impulsive dei *sistemi equivalenti* sono date da

$$h[n] = h_1[n] \otimes h_2[n] \qquad \text{(sistemi in } \textit{cascata}\text{)} \tag{6.1.15a}$$

$$h[n] = h_1[n] + h_2[n] \qquad \text{(sistemi in } \textit{parallelo}\text{)} \tag{6.1.15b}$$

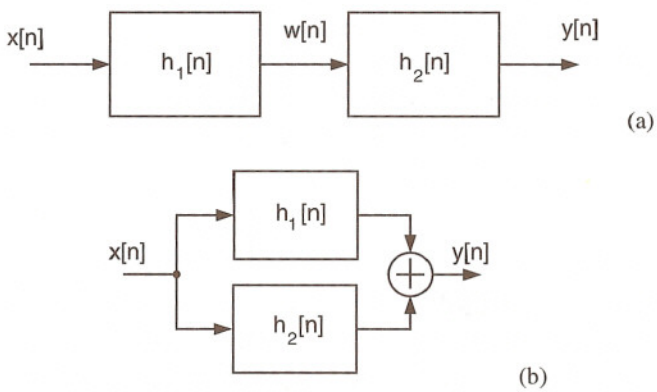


Figura 6.3 Sistemi lineari e stazionari in cascata (a) e in parallelo (b)

6.1.3 Risposta in frequenza di un SLS

La definizione di *risposta in frequenza* di un SLS a tempo discreto stabile non differisce apprezzabilmente da quella introdotta nel Paragrafo 4.2 cui si rimanda per ulteriori dettagli. Dunque diremo che:

i) la risposta in frequenza di un SLS a tempo discreto è la trasformata di Fourier della risposta impulsiva $h[n]$ del sistema stesso:

$$\overline{H}(f) = \sum_{n=-\infty}^{+\infty} h[n]\, e^{-j2\pi nfT} \tag{6.1.16}$$

ii) la risposta in frequenza $\overline{H}(f)$ è il rapporto fra le trasformate $\overline{Y}(f)$ e $\overline{X}(f)$

rispettivamente della sequenza di uscita $y[n]$ e di quella d'ingresso $x[n]$:

$$\overline{H}(f)=\frac{\overline{Y}(f)}{\overline{X}(f)} \tag{6.1.17}$$

iii) la risposta in frequenza $\overline{H}(f)$ è data dal rapporto fra la sequenza di uscita $y[n]$ e quella di ingresso $x[n]$ quando $x[n]$ è una oscillazione complessa alla frequenza f:

$$\overline{H}(f)=\left.\frac{y[n]}{x[n]}\right|_{x[n]=e^{j2\pi fT}} \tag{6.1.18}$$

La dimostrazione dell'equivalenza fra le diverse espressioni della risposta in frequenza di un SLS a tempo discreto è lasciata al lettore per esercizio. Data la risposta in frequenza $\overline{H}(f)$, definiamo anche per i sistemi a tempo discreto la *risposta in ampiezza*

$$\overline{A}(f)=|\overline{H}(f)| \tag{6.1.19}$$

che permette di stabilire le caratteristiche di *selettività* di un SLS, e la sua *risposta in fase*

$$\overline{\theta}(f)=\angle\overline{H}(f) \tag{6.1.20}$$

Esempio 6.2

Data la risposta in frequenza $\overline{H}(f)$ di un SLS, troviamo l'espressione del segnale di uscita $y[n]$ quando l'ingresso è la sequenza periodica di periodo N_0 $x[n]=x[n+N_0]$.

La sequenza data può essere scomposta in serie discreta (antitrasformata) di Fourier (5.5.2):

$$x[n]=\sum_{k=0}^{N_0-1}\overline{X}_k\, e^{j\frac{2\pi kn}{N_0}} \tag{E6.2.1}$$

cioè come *somma pesata* (combinazione lineare) di N_0 oscillazioni sinusoidali complesse a tempo discreto alle frequenze $f_k=k/(N_0T)$, $k=0,1,...,N_0-1$. Poiché il sistema è *lineare*, possiamo applicare il principio di sovrapposizione degli effetti, osservando che ciascuna di queste oscillazioni viene modificata dal sistema in ampiezza e fase in ragione del valore della risposta in frequenza in

corrispondenza della rispettiva frequenza di oscillazione. L'espressione del segnale d'uscita è dunque

$$y[n] = \sum_{k=0}^{N_0-1} \overline{X}_k \cdot \overline{H}\left(\frac{k}{N_0 T}\right) e^{j\frac{2\pi kn}{N_0}} \tag{E6.2.2}$$

cioè la trasformata discreta di Fourier $\overline{Y}_k$ del segnale d'uscita $y[n]$ è

$$\overline{Y}_k = \overline{X}_k \cdot \overline{H}\left(\frac{k}{N_0 T}\right) \tag{E6.2.3}$$

❑

La *condizione di non distorsione* già enunciata nel Paragrafo 4.3 per i segnali a tempo continuo viene riformulata come segue:

$$y[n] = K\, x[n - n_0] \tag{6.1.21}$$

ove K ed n_0 rappresentano rispettivamente il guadagno e il ritardo del sistema. Nel dominio della frequenza, la condizione (6.1.21) si traduce nei due seguenti requisiti rispettivamente per la risposta in ampiezza e la risposta in fase:

$$\overline{A}(f) = K \tag{6.1.22a}$$

$$\overline{\theta}(f) = -2\pi f n_0 T \tag{6.1.22b}$$

Come già precisato nel Paragrafo 4.3, è sufficiente che le condizioni (6.1.22) siano verificate *nell'ambito della banda del segnale* d'ingresso per garantire assenza di distorsioni.

6.1.4 Filtri a tempo discreto

Il concetto di SLS selettivo in frequenza, o *filtro*, viene esteso direttamente al caso dei sistemi a tempo discreto, con identica nomenclatura. Naturalmente, le caratteristiche di selettività di un filtro a tempo discreto con risposta in frequenza $\overline{H}(f)$ (che è una funzione periodica di periodo $1/T$) sono determinate dall'andamento della sua risposta in ampiezza $|\overline{H}(f)|$ in *un solo periodo della funzione*, ad esempio nell'intervallo $[-1/2T, 1/2T)$. Limitando l'analisi di $|\overline{H}(f)|$ a questo intervallo di frequenze occorre tener conto del fatto che le "basse frequenze" continuano a essere quelle prossime alla frequenza nulla, mentre le "alte fre-

quenze" sono quelle prossime al limite superiore dell'intervallo, cioè $1/2T$. Queste considerazioni giustificano la classificazione dei filtri ideali riportata in Figura 6.4 per la quale si possono ripetere le stesse osservazioni esposte nella classificazione dei sistemi analogici del Paragrafo 4.3, in particolare quelle riguardo la causalità.

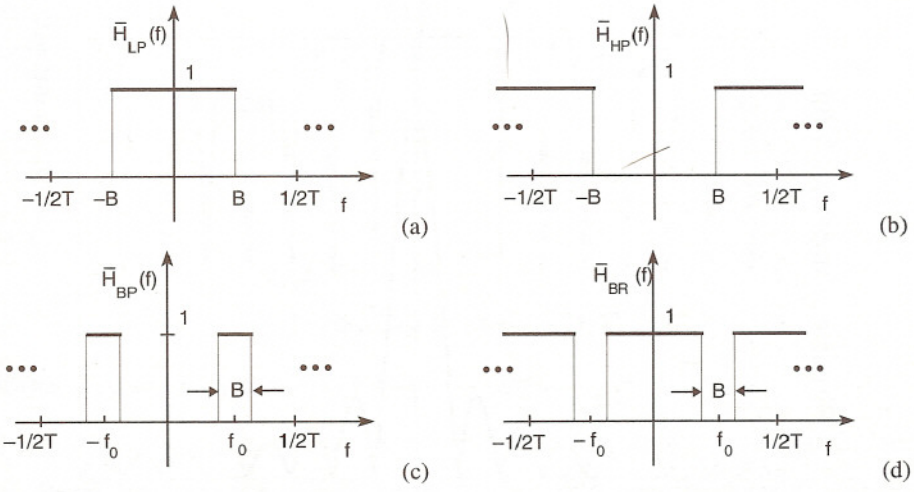


Figura 6.4 Filtri ideali a tempo discreto: passa-basso (a); passa-alto (b); passa-banda (c); elimina-banda (d)

Esempio 6.3

Riconsideriamo il *filtro a media mobile* dell'Esempio 6.1. La sua risposta impulsiva è

$$h[n]=\frac{1}{N}(u[n]-u[n-N]) \tag{E6.3.1}$$

e la sua risposta in frequenza è quindi (si veda l'Esempio 5.2)

$$\bar{H}(f)=\sum_{n=0}^{N-1}h[n]\,e^{-j2\pi nfT}=e^{-j\pi(N-1)fT}\,\frac{\sin(\pi NfT)}{N\sin(\pi fT)} \tag{E6.3.2}$$

La risposta in ampiezza del filtro è allora

$$\left|\bar{H}(f)\right|=\frac{1}{N}\left|\frac{\sin(\pi NfT)}{\sin(\pi fT)}\right| \tag{E6.3.3}$$

che è rappresentata in Figura 6.5, insieme alla relativa risposta in fase, nel caso

particolare $N = 16$. Si nota che $|\overline{H}(0)| = 1$, mentre le componenti ad alte frequenze (cioè in prossimità di $f = 1/2T$) sono sensibilmente attenuate rispetto a questo valore. Il filtro deve quindi considerarsi un *passa-basso*. Questa caratteristica è ulteriormente evidenziata dalla risposta in ampiezza *in dB* riportata in Figura 6.6.

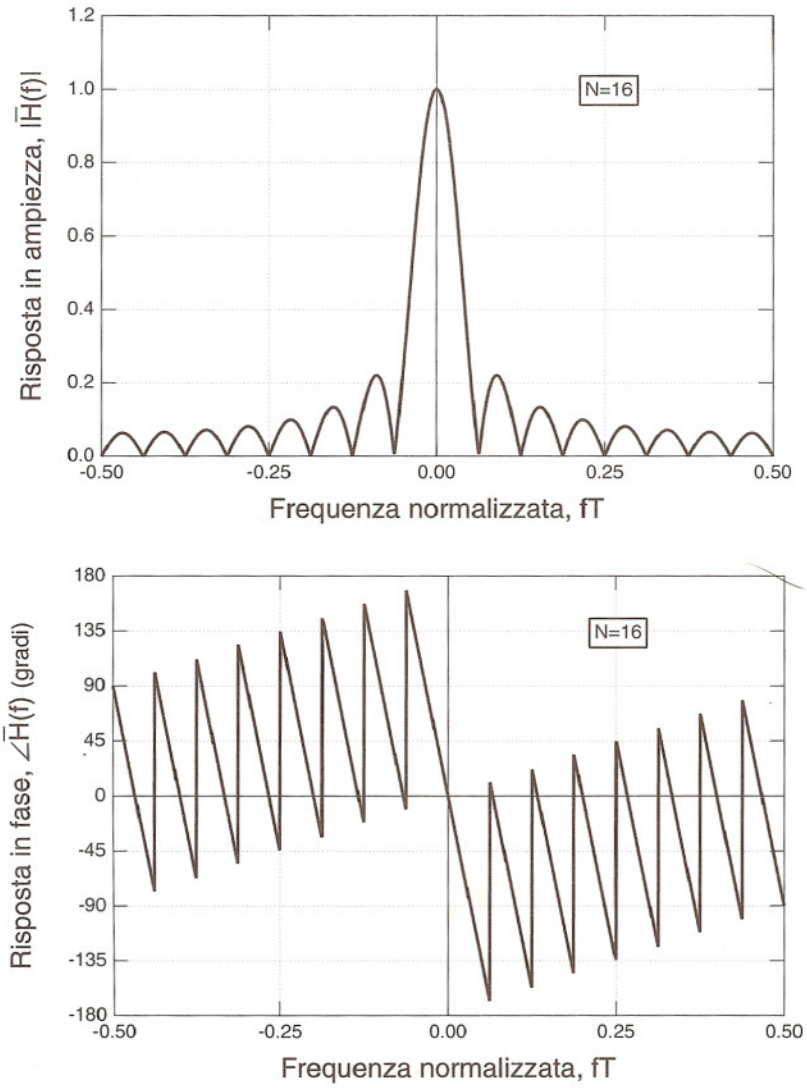


Figura 6.5 Risposta in ampiezza (a) e in fase (b) di un filtro a media mobile su 16 valori

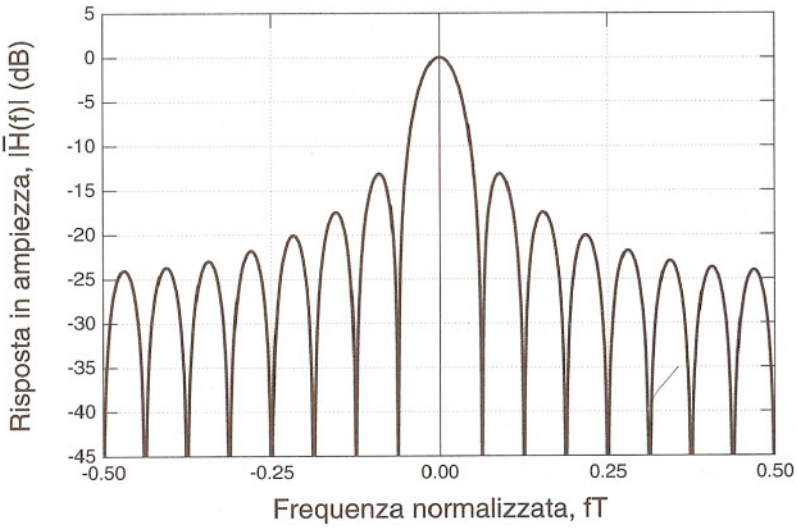


Figura 6.6 Risposta in ampiezza in dB del filtro di Figura 6.5

Il filtro a media mobile viene spesso utilizzato per eliminare le fluttuazioni rapide di un segnale (cioè per *smussare* il segnale) senza modificarne l'andamento generale di lungo termine. La Figura 6.7a mostra l'andamento dell'indice della borsa Italiana MIB durante l'anno 1998. La sequenza dei valori dell'indice MIB alla chiusura è molto frastagliata; per eliminare le fluttuazioni si può usare un filtro a media mobile con $N = 16$ (cioè quello appena considerato) ottenendo la curva *smussata* di Figura 6.7b. ❑

Esempio 6.4

Calcoliamo la risposta impulsiva dei filtri passa-basso e passa-alto ideale a tempo discreto di Figura 6.4a-b. La risposta in frequenza del passa-basso ideale è una funzione rect(·) *periodicizzata* con periodo frequenziale $1/T$:

$$\overline{H}_{LP}(f) = \frac{1}{T}\sum_{k=-\infty}^{\infty} T \cdot \text{rect}\left(\frac{f - k/T}{2B}\right) \qquad \text{(E6.4.1)}$$

per cui (si confronti anche con lo svolgimento dell'Esempio 5.7)

$$h_{LP}[n] = 2BT\,\text{sinc}(2Bt)\big|_{t=nT} = 2BT\,\text{sinc}(2nBT) \qquad \text{(E6.4.2)}$$

Per quanto riguarda il filtro passa-alto, si ha:

$$\overline{H}_{HP}(f) = 1 - \overline{H}_{LP}(f) \tag{E6.4.3}$$

e quindi

$$h_{HP}[n] = \delta[n] - h_{LP}[n] = \delta[n] - 2BT\,\mathrm{sinc}(2nBT) \tag{E6.4.4}$$

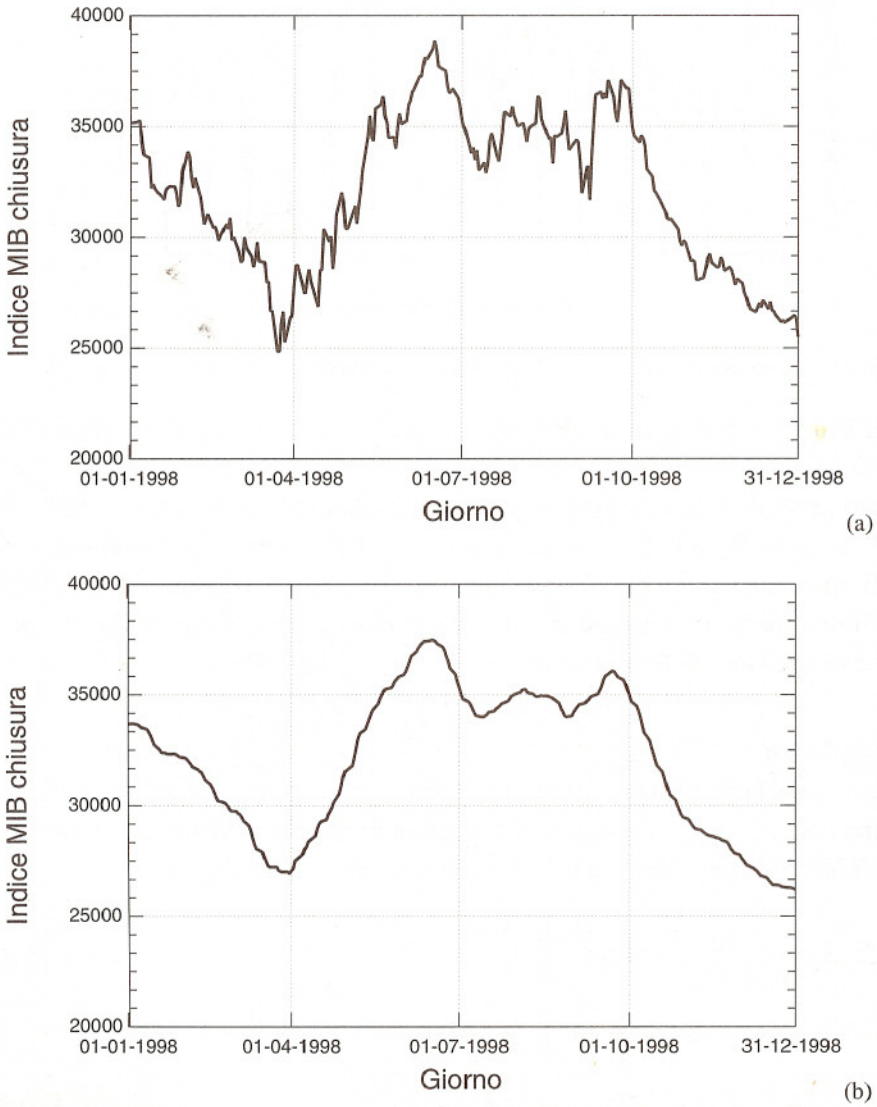


Figura 6.7 Andamento dell'indice MIB nel 1998 senza (a) e con (b) filtraggio a media mobile

Si noti che entrambe le risposte impulsive sono *non nulle* per $n < 0$. ❑

Dall'esempio precedente, e come già anticipato nella discussione generale, si nota che i filtri ideali *non* sono causali. Questa conclusione si può trarre anche dalla versione per segnali a tempo discreto del *criterio di Paley-Wiener*. Se la risposta in ampiezza del sistema è a quadrato sommabile sul periodo-base, cioè se

$$T \int_{-1/2T}^{1/2T} \left|\overline{H}(f)\right|^2 df < \infty \tag{6.1.23}$$

e verifica la condizione

$$\int_{-1/2T}^{1/2T} \left|\ln\left[\left|\overline{H}(f)\right|\right]\right| df < \infty \tag{6.1.24}$$

allora esiste una funzione reale $\overline{\theta}(f) = \overline{\theta}(f + 1/T)$ per cui

$$\overline{H}(f) = \left|\overline{H}(f)\right| e^{j\overline{\theta}(f)} \tag{6.1.25}$$

rappresenta la risposta in frequenza di un sistema a tempo discreto *causale*.

Esempio 6.5

Consideriamo il sistema di Figura 6.8a in cui la frequenza di campionamento è $f_c = 48$ kHz; il segnale di ingresso è inoltre

$$x(t) = \cos(2\pi f_0 t) \tag{E6.5.1}$$

con $f_0 = 16$ kHz. Determiniamo l'espressione del segnale d'uscita $y[n]$ sapendo che il *sistema nonlineare* (NL) a tempo discreto è senza memoria ed è caratterizzato dalla relazione ingresso-uscita $z[n] = x^2[n]$, e che la risposta impulsiva del successivo filtro è

$$h[n] = \frac{1}{2}\operatorname{sinc}\left(\frac{n}{2}\right) \tag{E6.5.2}$$

La sequenza in uscita al campionatore è

$$x[n] = \cos(2\pi n f_0 T) \tag{E6.5.3}$$

e, di conseguenza, la sequenza in uscita alla nonlinearità è data da

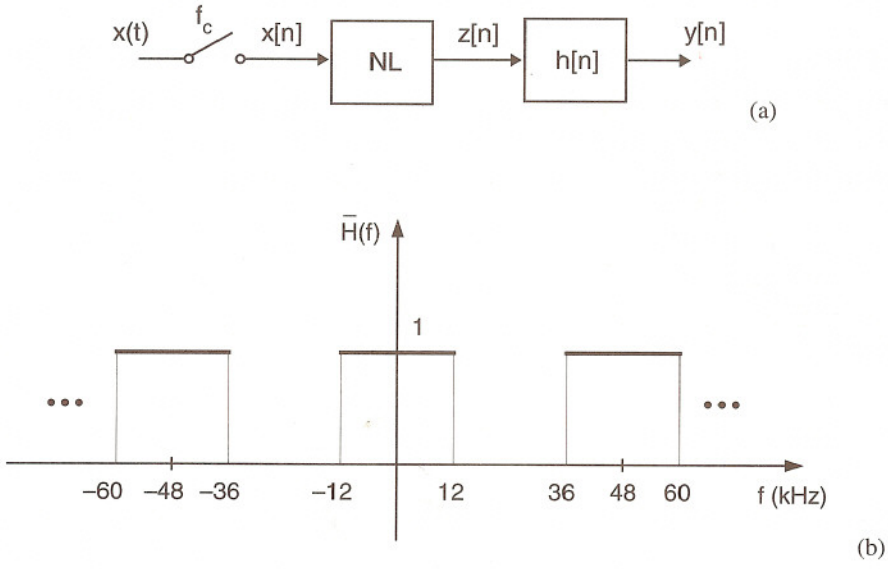


Figura 6.8 Sistema dell'Esempio 6.5 (a) e risposta in frequenza del relativo filtro $h[n]$ (b)

$$z[n] = \cos^2(2\pi n f_0 T) = \frac{1}{2}\left[1 + \cos(4\pi n f_0 T)\right] \tag{E6.5.4}$$

La sequenza $z[n]$ è costituita dalla somma di una costante (componente continua) e di una oscillazione cosinusoidale a frequenza $2f_0 = 32$ kHz. La sequenza $y[n]$ in uscita al filtro $h[n]$ è quindi

$$y[n] = \frac{1}{2}\overline{H}(0) + \frac{1}{2}\left|\overline{H}(2f_0)\right| \cos\left(4\pi n f_0 T + \angle\overline{H}(2f_0)\right) \tag{E6.5.5}$$

ove $\overline{H}(f) = \mathcal{F}[h[n]]$ è la risposta in frequenza del sistema a tempo discreto. Dall'Esempio 6.4 si trova che $h[n]$ è la risposta impulsiva di un filtro passa-basso ideale di banda $B = 1/4T = 12$ kHz la cui risposta in frequenza è rappresentata nella Figura 6.8b.

Si può verificare immediatamente che $\overline{H}(0) = 1$ e $\overline{H}(2f_0) = 0$, per cui la sequenza $y[n]$ è

$$y[n] = \frac{1}{2} \tag{E6.5.6}$$

❑

6.2 Cambiamento della frequenza di campionamento

Molto spesso, nella pratica dell'elaborazione numerica dei segnali è opportuno cambiare *in tempo reale* la cadenza con cui i campioni di una sequenza vengono elaborati. Se ad esempio si considera il campionamento di un segnale audio ad alta qualità, sappiamo che la frequenza di campionamento f_c deve essere maggiore di 40 kHz. Una frequenza standard per queste applicazioni è infatti la già citata $f_c = 48$ kHz. Supponiamo che il segnale campionato a questa velocità venga filtrato con un filtro numerico *passa-basso* ideale di banda $B = 15$ kHz, come in Figura 6.9. È chiaro che per questo nuovo segnale a banda ridotta la cadenza di campionamento di 48 kHz originaria risulta in qualche modo sovrabbondante. È conveniente allora introdurre un opportuno sistema per *cambiare* (in questo caso, ridurre) la cadenza di campionamento della sequenza $z[n]$ in uscita dal filtro, ad esempio passare da 48 a 32 kHz. La sequenza $y[m]$ in ingresso al convertitore D/A che chiude la catena di elaborazione può allora essere convertita con una frequenza $f_c' = 32$ kHz, usando un convertitore D/A a 32 kHz più semplice e meno costoso di uno alla frequenza originaria $f_c = 48$ kHz. Si noti che la variabile temporale delle sequenze $x[\cdot]$ e $y[\cdot]$ è stata indicata con due nomi diversi (rispettivamente n ed m), perché tali variabili devono intendersi riferite a due *diverse* frequenze di campionamento.

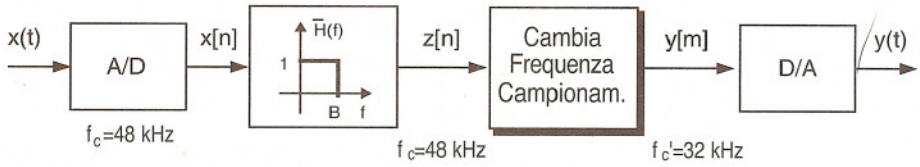


Figura 6.9 Cambiamento della frequenza di campionamento

Le operazioni principali di cambiamento della frequenza di campionamento f_c possono essere ricondotte a una combinazione di due funzioni base, il *sovracampionamento* e il *sottocampionamento*. Tali funzioni, opportunamente combinate, consentono di modificare f_c secondo un fattore *razionale* arbitrario, in modo da generare una nuova sequenza con cadenza $f_c' = (p/q) \cdot f_c$, p e q interi.

6.2.1 Sovracampionamento con interpolazione numerica

L'*aumento* della cadenza di campionamento secondo un fattore intero M è chiamato *sovracampionamento* ed è rappresentato dal blocco illustrato nella Figura 6.10a. Nell'operazione di sovracampionamento, anche quando non esplicitamente indicata, è sempre compresa una funzione di *interpolazione numerica*. Per

capire la necessità di quest'ultima funzione scomponiamo il blocco sovracampionatore nei due componenti elementari della Figura 6.10b. Il blocco ZP è quello che materialmente effettua l'aumento nella velocità di campionamento. Dalla sequenza di ingresso $x[n]$, avente frequenza di campionamento f_c, viene infatti prodotta la sequenza di uscita $x_{ZP}[m]$ avente frequenza $f_c' = M \cdot f_c$, ovvero periodo di campionamento $T' = T / M$. La generazione di questa nuova sequenza viene effettuata mediante un'operazione di *zero-padding*, cioè di riempimento con zeri, come esemplificato nella Figura 6.11, per $M = 4$. L'aumento della frequenza di campionamento viene cioè ottenuto mediante *inserimento di* $M-1$ *campioni nulli* tra ciascuna coppia di campioni consecutivi nella sequenza originaria $x[n]$:

$$x_{ZP}[m] = \begin{cases} x[n] & m = n \cdot M \\ \\ 0 & \text{altrimenti} \end{cases} \tag{6.2.1}$$

o, equivalentemente,

$$x_{ZP}[m] = \sum_{n=-\infty}^{+\infty} x[n]\delta[m - Mn] \tag{6.2.2}$$

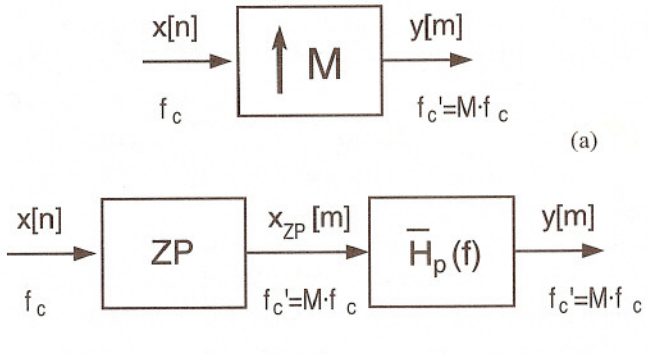


Figura 6.10 Rappresentazione simbolica (a) e realizzazione (b) del sovracampionatore con interpolazione

È interessante calcolare la trasformata di Fourier della sequenza sovracampionata, che va ovviamente riferita alla nuova frequenza di campionamento $f_c' = M \cdot f_c$:

$$\overline{X}_{ZP}(f) = \sum_{m=-\infty}^{+\infty} x_{ZP}[m]\, e^{-j2\pi mfT'} \tag{6.2.3}$$

Se si osserva che $x_{ZP}[m]$ è diversa da zero solo per $m = M \cdot n$ (nel qual caso $x_{ZP}[m] = x[n]$) e che $T' = T/M$, la (6.2.3) diventa

$$\overline{X}_{ZP}(f) = \sum_{n=-\infty}^{+\infty} x[n] e^{-j2\pi MnfT/M} = \sum_{n=-\infty}^{+\infty} x[n] e^{-j2\pi nfT} = \overline{X}(f) \tag{6.2.4}$$

cioè le trasformate delle sequenze $x[n]$ e $x_{ZP}[m]$ sono identiche!

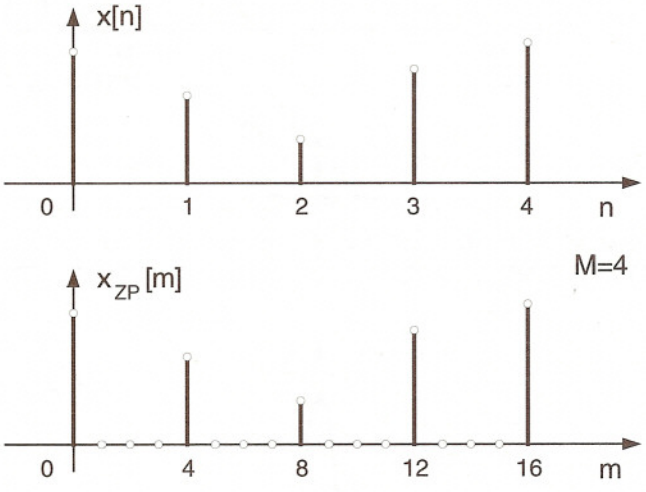


Figura 6.11 Riempimento con zeri e sovracampionamento

Questo risultato è solo apparentemente paradossale. Infatti le trasformate $\overline{X}_{ZP}(f)$ e $\overline{X}(f)$ sono uguali per ogni valore della frequenza, ma presentano una differenza fondamentale, evidenziata nella Figura 6.12 per $M = 4$: il periodo-base di $\overline{X}_{ZP}(f)$ è l'intervallo $[-1/2T', 1/2T']$ che, rispetto al corrispondente periodo-base $[-1/2T, 1/2T]$ di $\overline{X}(f)$, è *M volte più grande*. Ciò significa che elaborando $x_{ZP}[m]$ con un filtro *numerico*, cioè proprio con il secondo blocco dello schema di Figura 6.10b, si possono modificare *tutte le componenti frequenziali appartenenti all'intervallo* $[-1/2T', 1/2T']$. Se invece elaboriamo direttamente con un filtro numerico la sequenza originaria $x[n]$ possiamo agire soltanto sulle sue componenti frequenziali relative all'intervallo $[-1/2T, 1/2T]$.

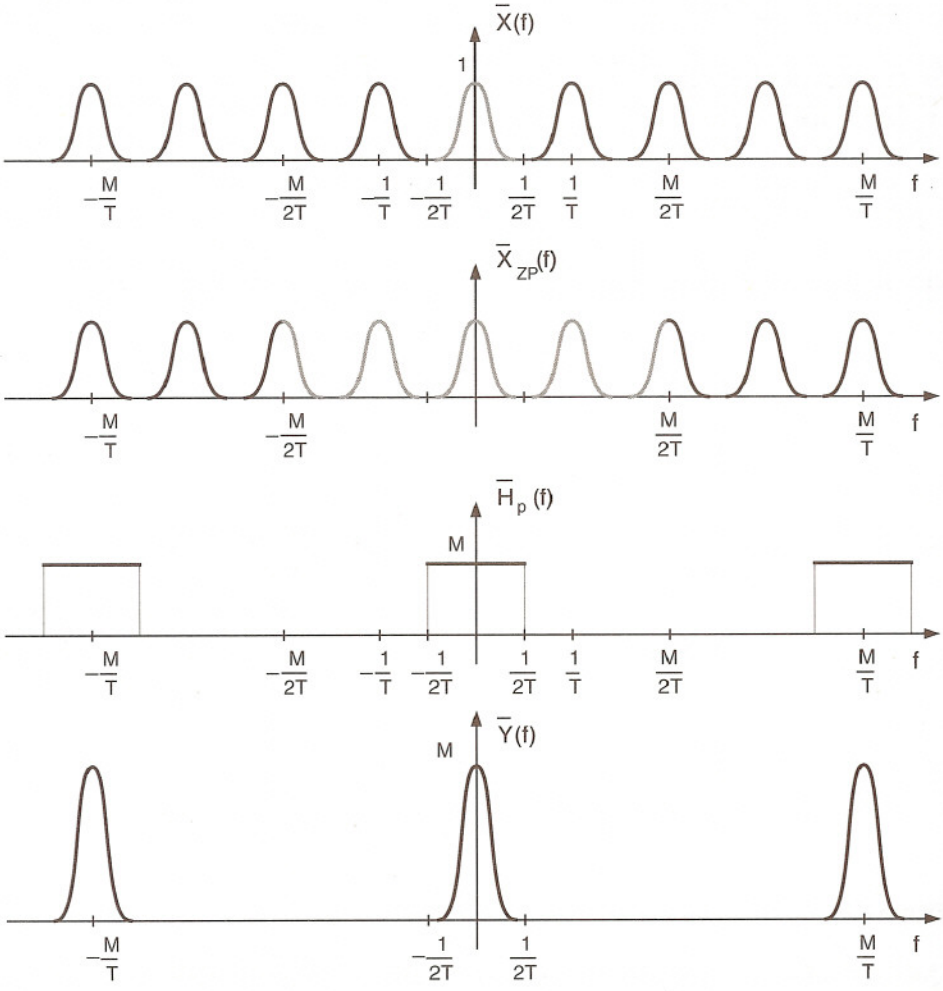


Figura 6.12 Trasformate della sequenza $x[m]$ d'ingresso (a) e della sequenza riempita con zeri $x_{ZP}[m]$ (b); risposta in frequenza del filtro interpolatore (c) e trasformata della sequenza sovracampionata d'uscita $y[m]$ (d).

Al blocco ZP della Figura 6.10b si fa quindi seguire un *filtro passa-basso ideale* $\overline{H}_p(f)$, avente banda $B = 1/2T = 1/(2MT')$ e guadagno in continua $\overline{H}_p(0) = M$. Questo filtro, la cui risposta in frequenza è illustrata nella Figura 6.12c, opera alla frequenza di campionamento f_c' e ha una funzione ben precisa, che risulta chiara esaminando l'andamento della trasformata $\overline{Y}(f)$ della sequenza di uscita $y[m]$ mostrato nella Figura 6.12d. Si nota che il filtro ha *cancellato* alcune

immagini viciniori presenti nello spettro $\overline{X}_{ZP}(f)$. Questo spettro è identico a quello che si sarebbe ottenuto campionando direttamente il segnale analogico $x(t)$ di partenza con la frequenza di campionamento f_c'!

Il filtro passa-basso ideale è chiamato *filtro interpolatore*. Non si deve confondere questo filtro *numerico* passa-basso con l'*interpolatore,* introdotto nel Paragrafo 5.1, che permette di ricostruire un segnale analogico a partire da una sequenza di campioni. Il nome di filtro interpolatore è, tuttavia, di uso comune per le ragioni seguenti. Se indichiamo con $h_p[m]$ la risposta impulsiva del filtro $\overline{H}_p(f)$, l'espressione del segnale sovracampionato $y[m]$ è (vedi la (6.2.2))

$$y[m] = x_{ZP}[m] \otimes h_p[m] = \sum_{n=-\infty}^{+\infty} x[n]\delta[m - Mn] \otimes h_p[m]$$

$$= \sum_{n=-\infty}^{+\infty} x[n]\, h_p[m - Mn] \tag{6.2.5}$$

Questa relazione è formalmente analoga alla (5.4.9) che riassume l'elaborazione del segnale compiuta dall'interpolatore propriamente detto.

La funzione di interpolatore svolta dal filtro digitale diviene ancora più chiara se si osserva che la risposta impulsiva $h_p[m]$ del filtro passa-basso ideale avente banda $1/(2MT')$ e guadagno M è

$$h_p[m] = \text{sinc}(m/M) \tag{6.2.6}$$

Sostituendo allora la (6.2.6) nella (6.2.5) si trova

$$y[m] = \sum_{n=-\infty}^{+\infty} x[n]\text{sinc}\left(\frac{m - Mn}{M}\right) \tag{6.2.7}$$

che ricorda la formula di interpolazione cardinale (5.4.21). La sequenza $y[m]$ è, quindi, una versione *interpolata* del segnale $x_{ZP}[m]$. In tale versione i campioni nulli deliberatamente introdotti vengono sostituiti da campioni interpolati che assicurano una maggiore "regolarità" del segnale sovracampionato, come illustrato nella Figura 6.13.

Esempio 6.6

Il segnale

$$x(t) = \text{sinc}^2\left(\frac{t}{2T}\right) \tag{E6.6.1}$$

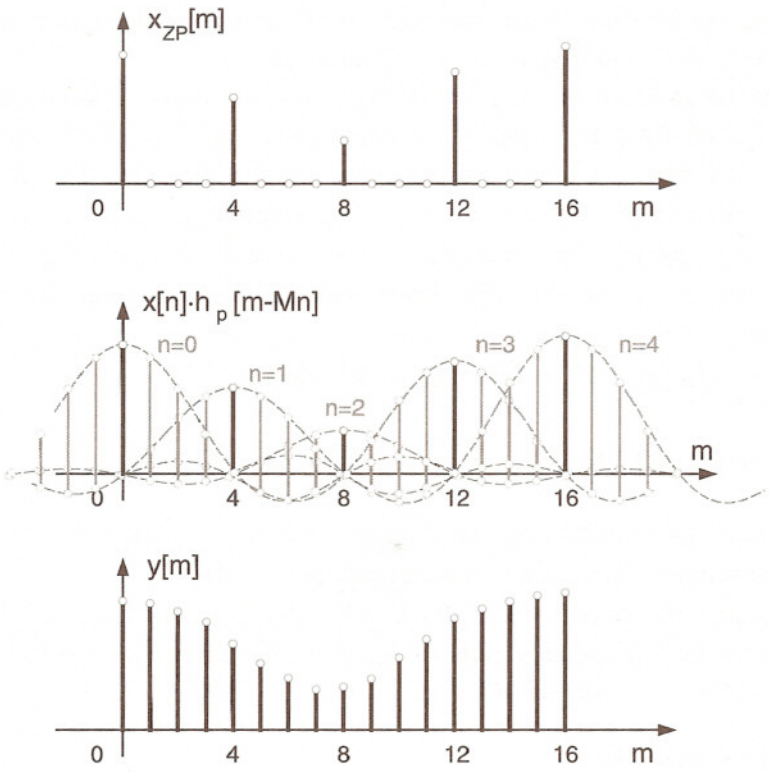


Figura 6.13 Interpolazione numerica

viene elaborato dai due sistemi illustrati in Figura 6.14, in entrambi i quali la frequenza di campionamento iniziale è $f_s = 1/T$. Le risposte impulsive dei filtri sono

$$h(t) = \frac{4}{T}\text{sinc}\left(\frac{2t}{T}\right), \quad h_1(t) = \frac{1}{T}\text{sinc}\left(\frac{t}{T}\right) \tag{E6.6.2}$$

Determiniamo e rappresentiamo gli spettri dei segnali $z(t)$ e $z_1(t)$ e stabiliamo quale di questi ultimi presenta minore distorsione rispetto al segnale $x(t)$.

Cominciamo col considerare la trasformata di Fourier del segnale $x(t)$:

$$X(f) = 2T(1 - 2|f|T)\text{rect}(fT) \tag{E6.6.3}$$

la trasformata $\overline{X}(f)$ della sequenza $x[n]$ ottenuta per campionamento sarà dunque

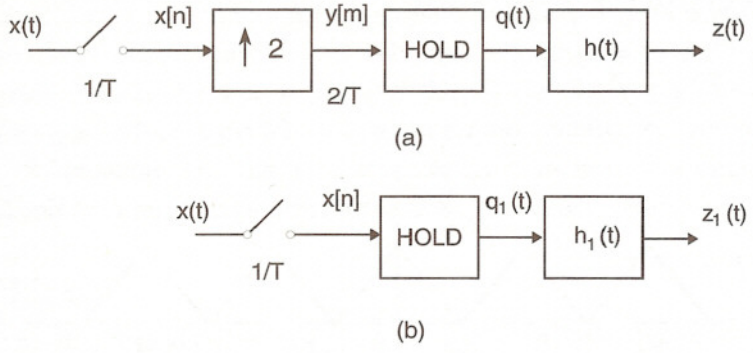


Figura 6.14 - Esempio di sovracampionamento

$$\overline{X}(f)=\frac{1}{T}\sum_{k=-\infty}^{+\infty}\overline{X}\left(f-\frac{k}{T}\right) \tag{E6.6.4}$$

il cui andamento è illustrato nella Figura 6.15a. La teoria del sovracampionamento appena sviluppata ci dice che il dispositivo di sovracampionamento della Figura 6.14 cancella dallo spettro di $x[n]$ tutte le immagini di $X(f)$ poste a cavallo dei *multipli dispari* della frequenza $1/T$. La trasformata $\overline{Y}(f)$ della sequenza $y[m]$ sarà quella illustrata in Figura 6.15b la cui espressione è

$$\overline{Y}(f)=\frac{1}{T}\sum_{k=-\infty}^{+\infty}\overline{X}\left(f-\frac{2k}{T}\right) \tag{E6.6.5}$$

A questo punto è immediato ricavare la trasformata del segnale interpolato $q(t)$:

$$Q(f)=\overline{Y}(f)\frac{T}{2}\operatorname{sinc}\left(\frac{fT}{2}\right)e^{-j\pi fT/2} \tag{E6.6.6}$$

ove si è tenuto conto che l'impulso di interpolazione è rettangolare di durata pari a $T/2$ (si ricordi che questo interpolatore opera a frequenza di campionamento $2/T$). Inoltre, il filtro della Figura 6.14a ha risposta in frequenza

$$H(f)=2\,\mathrm{rect}\left(\frac{f}{2/T}\right) \tag{E6.6.7}$$

e cancella *tutte* le immagini di $\overline{Y}(f)$ (vedi la (E6.6.6)). Lo spettro del segnale $z(t)$ all'uscita del sistema della Figura 6.14a è allora

$$Z(f) = Q(f)H(f) = \frac{1}{T} X(f)\, T \operatorname{sinc}\left(\frac{fT}{2}\right) e^{-j\pi fT/2}$$

$$= X(f) \operatorname{sinc}\left(\frac{fT}{2}\right) e^{-j\pi fT/2} \qquad \text{(E6.6.8)}$$

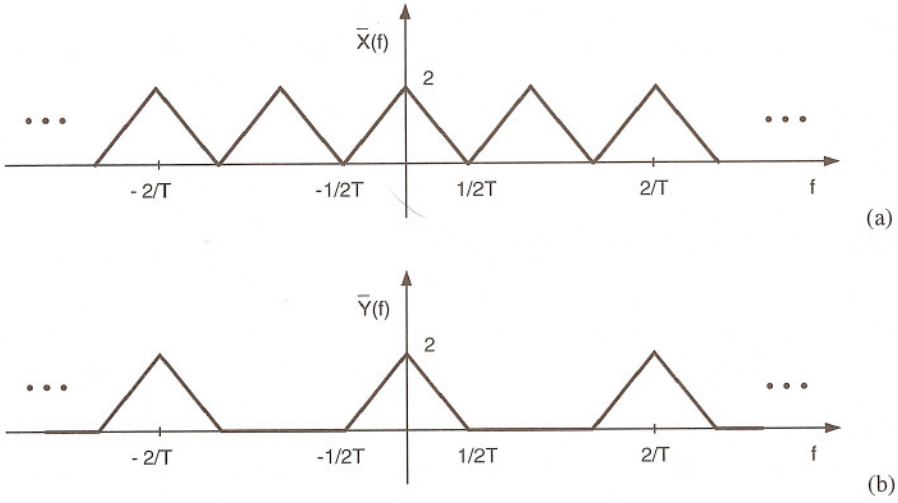


Figura 6.15 Spettri delle sequenze $x[n]$ (a) e $y[m]$ (b) dell'Esempio 6.6

Dunque, il sistema della Figura 6.14a introduce una *distorsione di ampiezza* sul segnale $x(t)$.

La trasformata del segnale interpolato $q_1(t)$ è invece

$$Q_1(f) = \overline{X}(f) T \operatorname{sinc}(fT) e^{-j\pi fT} \qquad \text{(E6.6.9)}$$

poiché stavolta l'impulso-base dell'interpolatore a mantenimento è un impulso rettangolare avente durata T. Il filtro della Figura 6.14b ha risposta in frequenza

$$H_1(f) = \operatorname{rect}(fT) \qquad \text{(E6.6.10)}$$

e cancella i contributi di tutte le immagini in $\overline{X}(f)$. Lo spettro del segnale $z_1(t)$ all'uscita del sistema della Figura 6.14b è allora

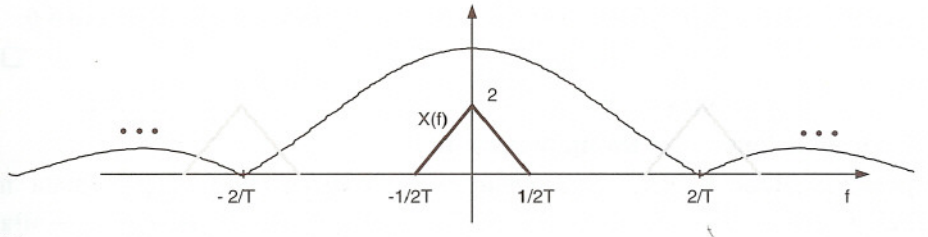
$$Z_1(f) = Q_1(f)H_1(f) = \frac{1}{T} X(f)\, T \operatorname{sinc}(fT) e^{-j\pi fT} = X(f) \operatorname{sinc}(fT) e^{-j\pi fT} \qquad \text{(E6.6.11)}$$

In analogia con quanto accade con il primo sistema, anche il secondo sistema introduce solo una distorsione di ampiezza. Se però si considera la Figura 6.16,

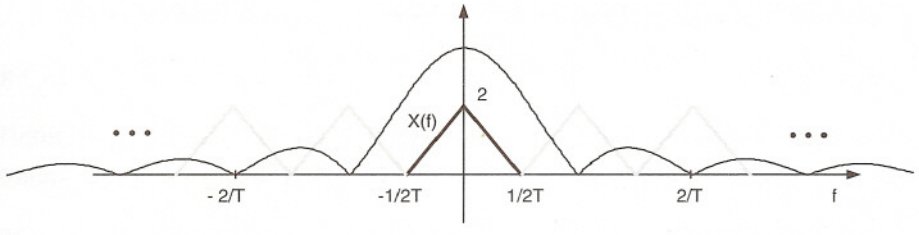
che rappresenta le due situazioni a confronto, si nota che la distorsione sperimentata nel sistema di Figura 6.14a (e rappresentata in Figura 6.16a) è evidentemente *minore* che nel caso della 6.14b poiché la trasformata dell'impulso dell'interpolatore ha un lobo principale più "largo" che distorce in misura minore. La minore distorsione è evidentemente un effetto benefico del *sovracampionamento*. Chiaramente, con fattori di sovracampionamento ancora più grandi sarebbe possibile ridurre maggiormente l'effetto distorcente dell'interpolatore. ❑

Esempio 6.7

Un lettore di Compact Disc (CD) ricostruisce il segnale analogico da inviare all'amplificatore audio di potenza a partire dalla sequenza di campioni $x[n]$ che è stata riletta dal disco medesimo. Come sappiamo, il campionamento in fase di registrazione (si veda l'Esempio 1.1) avviene alla frequenza $f_c = 44.1$ kHz in modo da rispettare con un certo margine la condizione di Nyquist relativamente alla banda del segnale $B = 20$ kHz. Molti dei lettori comunemente in commercio usano un circuito di *sovracampionamento e interpolazione* per meglio effettuare la conversione del segnale da numerico ad analogico.



(a)



(b)

Figura 6.16 Costruzione degli spettri di ampiezza dei segnali $z(t)$ (a) e $z_1(t)$ (b) dell'Esempio 6.6

Se si utilizzasse direttamente, per l'operazione di conversione, un convertitore D/A (e, quindi, un interpolatore a mantenimento) operante a 44.1 kHz, si verifi-

cherebbe la situazione rappresentata nella Figura 6.17a. Le immagini presenti nello spettro del segnale audio numerico sarebbero molto vicine alle componenti utili dello spettro e il filtro *analogico* anti-immagine $H_{AI}(f)$, presente necessariamente all'uscita del D/A (si ricordi la Figura 5.22), dovrebbe avere una risposta in ampiezza con una transizione estremamente *ripida* nella banda $[B, 1/T - B]$, cioè 20 kHz-24.1 kHz per ben cancellare le immagini. Sfortunatamente, un filtro con queste caratteristiche di selettività (ad esempio un filtro passivo con molti poli) presenta anche una risposta in fase estremamente non lineare non solo nella banda di transizione ma *anche in un intervallo di frequenze utili* vicine 20 kHz. Questo comporterebbe un'apprezzabile distorsione di fase del segnale e un'inaccettabile perdita di qualità per l'ascoltatore.

Usando un circuito di sovracampionamento e interpolazione numerica (fattori tipici di sovracampionamento sono $M = 4$ o 8) si possono invece già *allontanare (reiettare) le prime immagini* con l'interpolazione digitale, come illustrato nella Figura 6.17b. Il filtraggio anti-immagine analogico è quindi molto semplificato, nel senso che la zona di transizione dalla banda utile alla banda soppressa nella risposta in ampiezza del filtro è molto più ampia e le distorsioni di fase sono molto più limitate. Inoltre, la distorsione di ampiezza dovuta al convertitore D/A è pure ridotta, secondo il meccanismo visto nell'Esempio 6.6.

❑

6.2.2 Decimazione o sottocampionamento

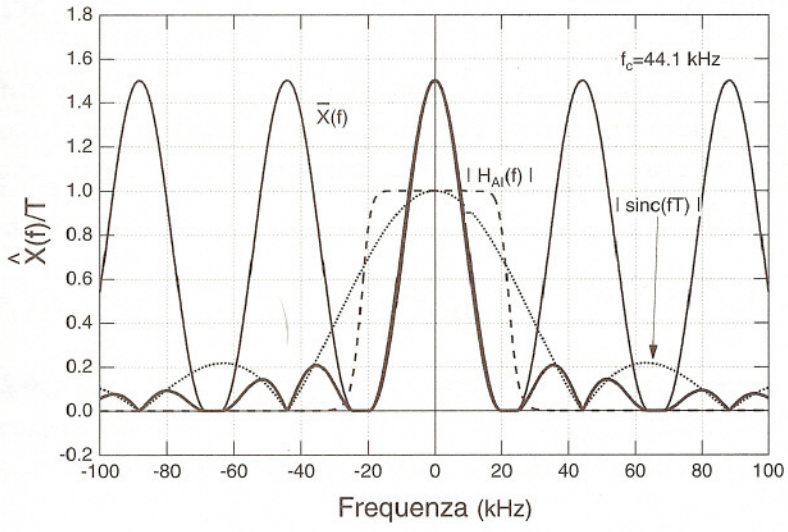
L'operazione duale del sovracampionamento è la *decimazione* rappresentata in Figura 6.18a-b. La sequenza in uscita $y[m]$ al decimatore è stavolta caratterizzata da una frequenza di campionamento $f_c' = 1/T' = f_c / M$ ed è ottenuta dalla sequenza di ingresso $x[n]$, a frequenza $f_c = 1/T$, come segue:

$$y[m] = x[M \cdot m] \tag{6.2.8}$$

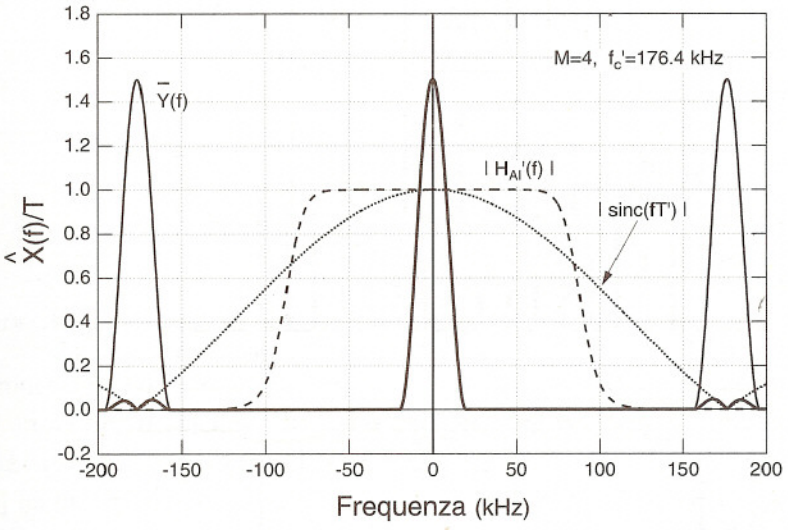
In pratica, si seleziona un campione di $x[n]$ ogni M (si "decimano" i campioni), e necessariamente si aumenta di un fattore M l'intervallo di campionamento: $T' = MT$.

La trasformata $\overline{Y}(f)$ della sequenza decimata sarà dunque

$$\overline{Y}(f) = \sum_{m=-\infty}^{+\infty} y[m]\, e^{-j2\pi mfT'} = \sum_{m=-\infty}^{+\infty} x[Mm]\, e^{-j2\pi mfMT} \tag{6.2.9}$$



(a)



(b)

Figura 6.17 Filtraggio anti-immagine senza (a) e con (b) sovracampionamento e interpolazione

D'altronde, il campione $x[Mm]$ può essere inteso come integrale di Fourier secondo la formula di antitrasformazione (5.2.7):

$$\overline{Y}(f) = \sum_{m=-\infty}^{+\infty} \left(T \int_{-1/2T}^{1/2T} \overline{X}(\nu)\, e^{j2\pi MmvT} d\nu \right) e^{-j2\pi mfMT}$$

$$= \int_{-1/2T}^{1/2T} \overline{X}(\nu) T \sum_{m=-\infty}^{+\infty} e^{-j2\pi Mm(f-\nu)T} d\nu \tag{6.2.10}$$

ove l'antitrasformazione usa come di consueto la variabile "muta" ν distinta dalla variabile f della trasformata. La serie di funzioni esponenziali a secondo membro può ora essere riscritta attraverso una formula di Poisson come un *pettine di funzioni* δ:

$$\begin{aligned} \overline{Y}(f) &= \int_{-1/2T}^{1/2T} \overline{X}(\nu) \sum_{k=-\infty}^{+\infty} \delta\left(f - \nu - \frac{k}{MT}\right) d\nu \\ &= \int_{f-1/2T}^{f+1/2T} \overline{X}(f-\alpha) \sum_{k=-\infty}^{+\infty} \delta\left(\alpha - \frac{k}{MT}\right) d\alpha \end{aligned} \tag{6.2.11}$$

avendo definito, per semplificare la formula, la variabile $\alpha = f - \nu$.

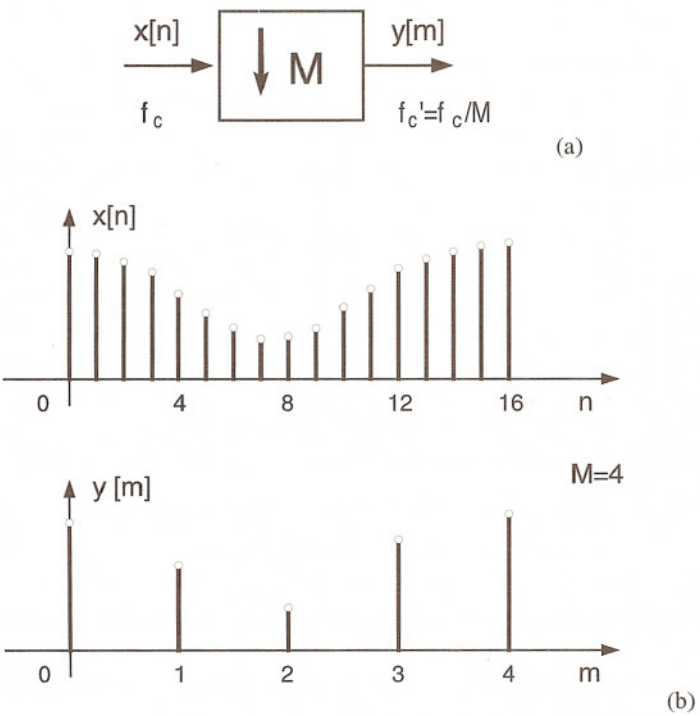


Figura 6.18 Simbolo (a) e funzione (b) di decimazione

Osserviamo che la funzione integranda della (6.2.11) è *periodica* nella variabile α di periodo $1/T$ e l'intervallo di integrazione ha ampiezza pari a un periodo.

Ne segue che il risultato dell'operazione di integrazione è *indipendente* dalla *posizione* del periodo di integrazione, cioè dal valore della variabile f. Scegliamo dunque per semplicità l'intervallo di integrazione $[0,1/T)$, e osserviamo che le uniche funzioni δ il cui punto di applicazione appartiene a questo intervallo sono quelle corrispondenti ai valori $k = 0,1,\ldots,M-1$ (si veda la Figura 6.19). Applicando la proprietà campionatrice della funzione δ si può quindi concludere che

$$\overline{Y}(f) = \sum_{k=0}^{M-1} \overline{X}\left(f - \frac{k}{MT}\right) = \sum_{k=0}^{M-1} \overline{X}(f - k \cdot f_c') \tag{6.2.12}$$

Si ottiene una *somma di M repliche* della trasformata della sequenza originaria, ciascuna traslata rispetto alla precedente di una quantità pari alla frequenza di campionamento "di uscita" $f_c' = 1/MT$.

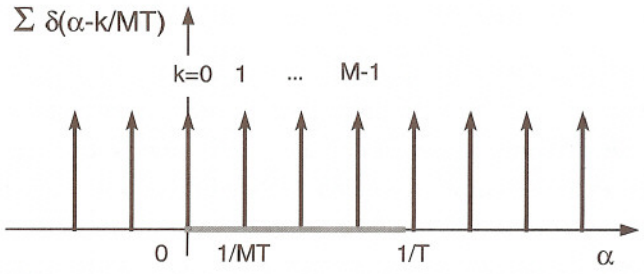


Figura 6.19 Integrazione del *pettine di* δ della relazione (6.2.11)

La conclusione di questo calcolo è che l'operazione di decimazione può comportare un fenomeno analogo a quello dell'*aliasing* nel campionamento di un segnale analogico (si confronti la (6.2.12) con la (5.4.7)). Infatti se la sequenza $x[n]$ ha una banda $B \leq 1/(2MT)$ (come nella Figura 6.20a) nello spettro $\overline{Y}(f)$ non si verifica alcuna sovrapposizione fra le repliche spettrali di $\overline{X}(f)$ (Figura 6.20b). Viceversa, se questa condizione (analoga alla condizione di Nyquist per il campionamento dei segnali analogici) non è verificata, si produce aliasing e lo spettro della sequenza decimata, nel suo proprio intervallo-base $[-f_c'/2, f_c'/2]$ è diverso da quello della sequenza originaria. Per questa ragione è importante che l'operazione di decimazione sia preceduta da un filtraggio numerico passa-basso anti-aliasing che limiti la banda della sequenza $x[n]$ a un valore minore del limite $1/(2MT)$.

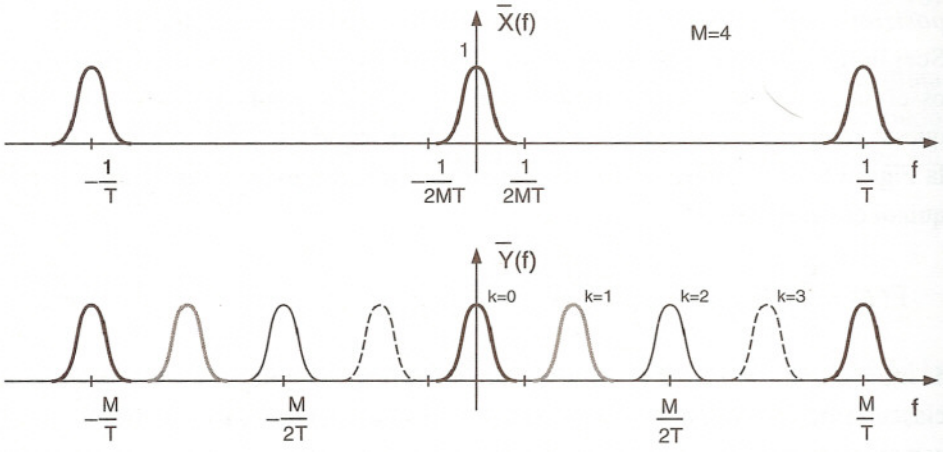


Figura 6.20 Decimazione senza aliasing

Esempio 6.8

Descriviamo adesso uno schema per la conversione alla nuova frequenza di campionamento $f_c' = 32$ kHz di un segnale audio campionato a frequenza $f_c = 48$ kHz. Osserviamo preliminarmente che le due frequenze stanno in un rapporto *razionale*, cosicché il problema è risolubile con le tecniche appena viste. In particolare, poiché $f_c'/f_c = 2/3$, è necessario disporre in cascata un sovracampionatore/interpolatore di un fattore 2 (che innalza la frequenza di campionamento a 96 kHz) seguito da un decimatore di un fattore 3 (che la riduce ai 32 kHz richiesti), come indicato nella Figura 6.21a. La realizzazione del sistema richiede, per quanto detto nei paragrafi precedenti, l'impiego di un filtro interpolatore (nel sovracampionatore) e di un filtro anti-aliasing (nel decimatore). Sia il primo che il secondo filtro devono essere dei passa-basso ideali aventi rispettivamente banda $B_1 = 20$ kHz e banda $B_2 = 15$ kHz. Poiché i due si trovano direttamente in cascata, il primo dei due (avente banda più larga) risulta evidentemente inutile. La disposizione finale del sistema è quindi quella di Figura 6.21b, in cui

$$h[p] = \frac{5}{8}\text{sinc}\left(\frac{5p}{16}\right) \qquad \text{(E6.8.1)}$$

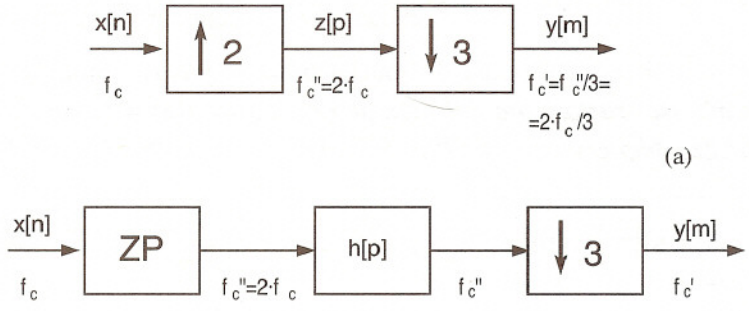


Figura 6.21 Cambiamento della frequenza di campionamento da 48 a 32 kHz

❑

6.3 Cenni alla trasformata Z di una sequenza

6.3.1 Definizione di trasformata Z e zone di convergenza

Le sequenze a *energia illimitata* non possiedono in generale trasformata di Fourier in senso ordinario, poiché la serie (5.2.5) è usualmente non convergente. In stretta analogia a quanto rapidamente accennato nel Paragrafo 3.6 relativamente alla trasformata di Laplace per segnali a tempo continuo, si introduce in questi casi la sequenza

$$\tilde{x}[n] = x[n] \cdot \left(\frac{1}{r}\right)^n \quad , \quad r > 0 \tag{6.3.1}$$

che, contrariamente alla sequenza originaria, tende rapidamente a zero quando $n \to \infty$. Ad esempio, la sequenza $x[n] = u[n]$ non ammette trasformata di Fourier ordinaria, mentre la $\tilde{x}[n]$ definita come sopra è una sequenza esponenziale smorzata che ammette trasformata se $r > 1$.

Calcoliamo ora la trasformata della nuova sequenza "smorzata" $\tilde{x}[n]$:

$$\overline{\tilde{X}}(f) = \sum_{n=-\infty}^{\infty} \tilde{x}[n]\, e^{-j2\pi nfT} = \sum_{n=-\infty}^{\infty} x[n]\, r^{-n} e^{-j2\pi nfT} = \sum_{n=-\infty}^{\infty} x[n]\, (r e^{j2\pi fT})^{-n} \tag{6.3.2}$$

In analogia con il procedimento che porta dalla trasformata di Fourier a quella di Laplace, possiamo anche qui definire un'*unica* variabile complessa espressa in forma polare come segue:

$$z \triangleq r e^{j2\pi fT} \quad , \quad r = |z| \geq 0 \quad , \quad -\pi < \angle z \leq \pi \quad , \Rightarrow \quad -\frac{1}{2T} \leq f < \frac{1}{2T} \tag{6.3.3}$$

e interpretare poi la trasformata di Fourier della sequenza "modificata" $\tilde{x}[n]$ come una *diversa trasformata del segnale originario* $x[n]$ dipendente appunto dalla variabile *complessa* $z = re^{j2\pi fT}$:

$$\boxed{X(z) = \mathcal{Z}[x[n]] \triangleq \sum_{n=-\infty}^{+\infty} x[n]\, z^{-n}} \tag{6.3.4}$$

Questa relazione definisce la *trasformata* $\mathcal{Z}$ bilatera della sequenza $x[n]$. La differenza principale rispetto alla trasformata di Fourier sta evidentemente nel fatto che questa trasformata esiste *in senso ordinario* anche in molti casi in cui la trasformata di Fourier della sequenza originaria non esiste. Per il segnale gradino abbiamo infatti

$$X(z) = \bar{\tilde{X}}(f) = \sum_{n=-\infty}^{\infty} \tilde{x}[n]\, e^{-j2\pi nfT} = \sum_{n=0}^{\infty} u[n]\, r^{-n} e^{-j2\pi nfT} = \sum_{n=0}^{\infty} \left(r^{-1} e^{-j2\pi fT}\right)^n$$

$$= \frac{1}{1 - r^{-1} e^{-j2\pi fT}} = \frac{1}{1 - z^{-1}} \quad , \quad r = |z| > 1 \tag{6.3.5}$$

È chiaro dunque che la sequenza gradino unitario ammette trasformata $\mathcal{Z}$ ordinaria e che questa è $U(z) = 1/(1 - z^{-1})$. Altrettanto chiaro dalla breve discussione sul ruolo del "parametro di smorzamento" $r = |z|$ è che tale risultato deve intendersi valido solo per *particolari* valori di r, in questo caso $r > 0$. Se ad esempio consideriamo la sequenza $x[n] = a^n u[n]$, $a > 1$, il minimo valore di r che garantisce smorzamento sufficiente è stavolta $r = a$, e dunque la trasformata sarà valida solo se $|z| > a$. Una relazione di questo tipo identifica una *zona* del piano complesso della variabile z in cui la trasformata $\mathcal{Z}$ esiste, in cui cioè la serie nella (6.3.5) è convergente.

Dal punto di vista squisitamente matematico, l'espressione (6.3.4) della trasformata $\mathcal{Z}$ rappresenta una *serie di potenze* o *serie di Taylor-Laurent* di variabile complessa. Come abbiamo anticipato con considerazioni euristiche, la serie di Laurent converge in un dominio del piano complesso descritto dalla relazione

$$R_1 < |z| < R_2 \tag{6.3.6}$$

ove R_1 ed R_2 sono i cosiddetti *raggi di convergenza* e dipendono dalla particolare sequenza $x[n]$ di cui si cerca la trasformata. Può accadere in particolare che si abbia $R_1 = 0$ e/o $R_2 = +\infty$.

Esempio 6.9

Calcoliamo la trasformata $\mathcal{Z}$ della sequenza

$$x[n] = a^n u[n] \tag{E6.9.1}$$

rappresentata in Figura 6.22. Applicando la definizione (6.3.4) si ha

$$X(z) = \sum_{n=-\infty}^{+\infty} a^n u[n]\, z^{-n} = \sum_{n=0}^{+\infty} a^n\, z^{-n} = \sum_{n=0}^{+\infty} \left(az^{-1}\right)^n \tag{E6.9.2}$$

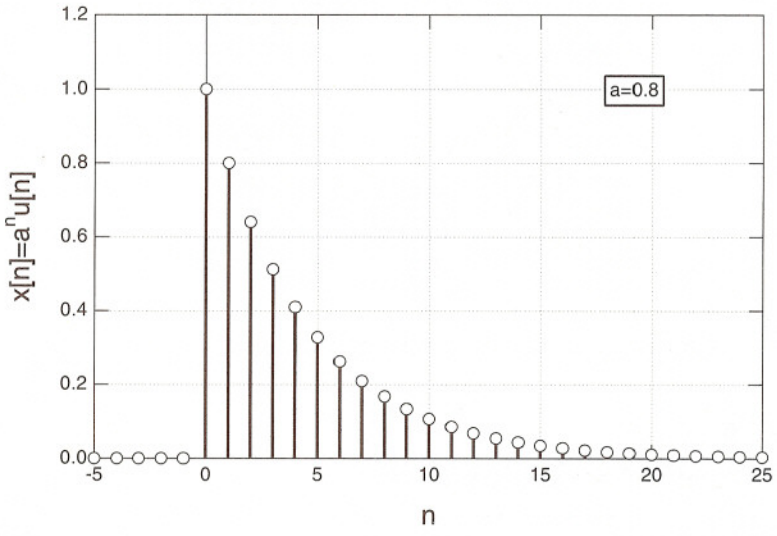


Figura 6.22 Sequenza dell'Esempio 6.9

La serie a secondo membro è una serie geometrica di ragione az^{-1}. Affinché questa converga è necessario che la sua ragione sia in modulo minore di 1, che si abbia cioè $|z| > |a|$. Sotto questa ipotesi la trasformata cercata è

$$X(z) = \frac{1}{1 - az^{-1}} \tag{E6.9.3}$$

La zona di convergenza $|z| > |a|$ di $X(z)$ è rappresentata in Figura 6.23 quando a è un numero reale positivo. Nel caso particolare $a = 1$, si ottiene la sequenza gradino unitario

$$x[n] = u[n] \tag{E6.9.4}$$

la cui trasformata $\mathcal{Z}$ è ricavabile come caso particolare dalla (E6.9.3):

$$X(z) = \frac{1}{1 - z^{-1}} \tag{E6.9.5}$$

con zona di convergenza $|z| > 1$, come già ricavato dalla discussione preliminare alla definizione di trasformata. ❑

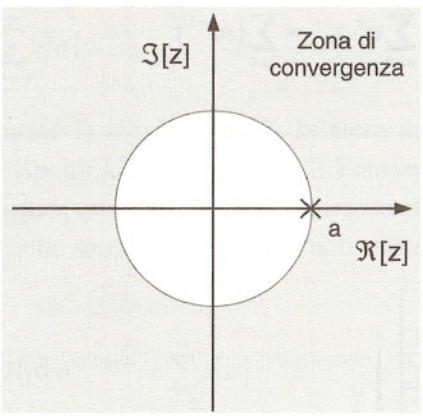


Figura 6.23 Zona di convergenza della trasformata dell'Esempio 6.9

Esempio 6.10

Calcoliamo la trasformata $\mathcal{Z}$ della sequenza

$$y[n] = -a^n u[-n-1] \tag{E6.10.1}$$

rappresentata in Figura 6.24. Applicando la definizione (6.3.4) si ha

$$Y(z) = -\sum_{n=-\infty}^{+\infty} a^n u[-n-1]\, z^{-n} = -\sum_{n=-\infty}^{-1} a^n\, z^{-n} = -\sum_{n=1}^{+\infty} \left(a^{-1} z\right)^n$$

$$= 1 - \sum_{n=0}^{+\infty} \left(a^{-1} z\right)^n \tag{E6.10.2}$$

Di nuovo, la serie a secondo membro della (E6.10.2) è una serie geometrica di ragione $a^{-1}z$. Se la ragione è in modulo minore di uno, cioè se $|z| < |a|$, la serie è convergente, e la trasformata cercata è

$$Y(z) = 1 - \frac{1}{1 - a^{-1} z} = \frac{-a^{-1} z}{1 - a^{-1} z} = \frac{1}{1 - a\, z^{-1}} \tag{E6.10.3}$$

cioè esattamente identica a quella della sequenza (completamente differente)

dell'Esempio 6.9! La zona di convergenza, rappresentata in Figura 6.25, è però $|z| < |a|$, cioè è esattamente *complementare* a quella dell'esempio precedente.

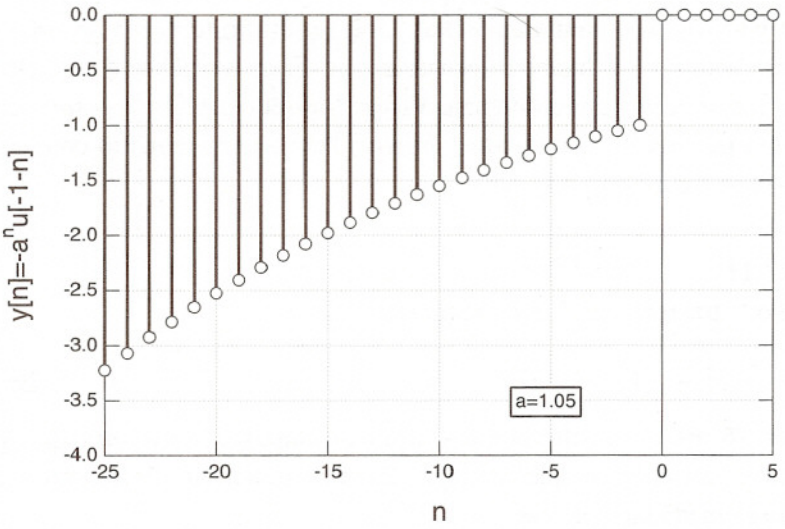


Figura 6.24 Sequenza dell'Esempio 6.10

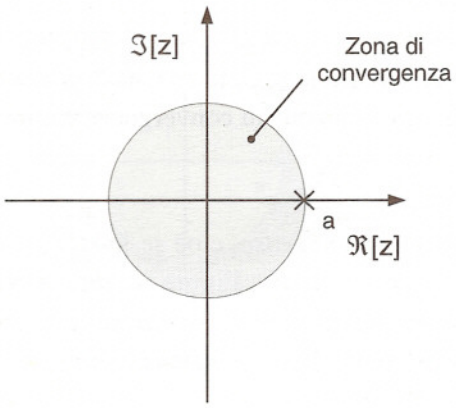


Figura 6.25 Zona di convergenza della trasformata dell'Esempio 6.10

❑

I due esempi appena visti indicano alcune proprietà generali della trasformata $\mathcal{Z}$ bilatera. Come è chiaro, assegnata una trasformata $X(z)$ non è in generale

possibile risalire univocamente alla sequenza $x[n]$ se non si specifica la zona di convergenza. Le sequenze dei due esempi sono infatti differenti, ma possiedono la stessa trasformata, ovviamente con differenti zone di convergenza. I due esempi sono stati scelti perché casi particolari di tendenze generali. Infatti, le sequenze *causali* (o comunque *destre*, cioè semiinfinite verso destra), come quella dell'Esempio 6.9, hanno in generale zone di convergenza date dalla parte di piano *esterna* a una circonferenza. Viceversa, sequenze *anticausali* (diverse da zero solo per $n<0$, o comunque *sinistre*) hanno come zona di convergenza un cerchio con centro nell'origine.

Esempio 6.11

Calcoliamo la trasformata $\mathcal{Z}$ della sequenza

$$w[n]=-b^n u[-n-1]+a^n u[n] \tag{E6.11.1}$$

di Figura 6.26. Essa è costituita dalla somma di una sequenza *anticausale* e di una *causale* analoghe a quelle degli Esempi 6.9-6.10. Possiamo dunque concludere immediatamente che

$$W(z)=\frac{1}{1-bz^{-1}}+\frac{1}{1-az^{-1}}=\frac{2-(a+b)z^{-1}}{\left(1-az^{-1}\right)\left(1-bz^{-1}\right)} \tag{E6.11.2}$$

nell'ipotesi che esista una zona di convergenza. Infatti la zona di convergenza è individuata dalla *intersezione* fra le regioni di convergenza della parte causale e di quella anticausale di $w[n]$. La zona di convergenza è allora

$$|a|<|z|<|b| \tag{E6.11.3}$$

ma solo se questa espressione ha senso, cioè se $|a|<|b|$, come in Figura 6.27, e come nell'esempio numerico della Figura 6.26; se viceversa $|a|\geq|b|$ la trasformata $\mathcal{Z}$ non esiste, perché non esiste nessuna zona del piano z in cui entrambi i termini di $w[n]$ danno luogo a serie convergenti. ❑

Così come succede per la trasformata di Laplace, si definisce anche una versione *monolatera* della trasformata $\mathcal{Z}$:

$$X(z)=\sum_{n=0}^{\infty}x[n]z^{-n} \tag{6.3.6}$$

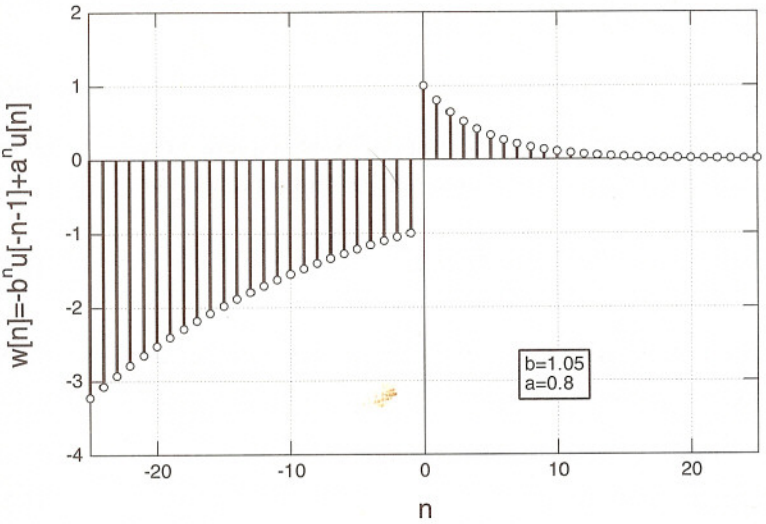


Figura 6.26 Sequenza dell'Esempio 6.11

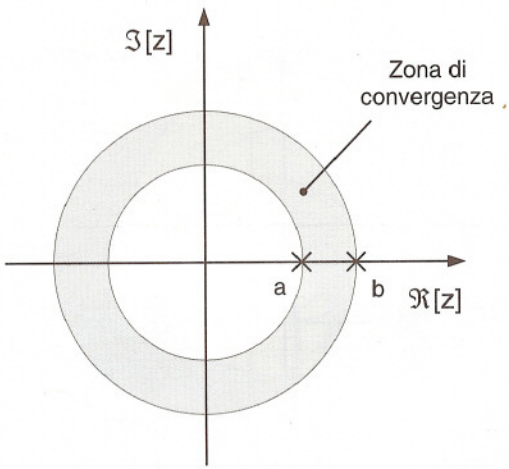


Figura 6.27 Zona di convergenza della trasformata dell'Esempio 6.11

che coincide con la trasformata bilatera già definita quando le sequenze sono causali. Nel caso di sequenze causali peraltro, la questione delle zone di convergenza si semplifica, come testimonia il seguente esempio.

Esempio 6.12
Calcoliamo la trasformata $\mathcal{Z}$ della sequenza *causale*

$$c[n] = a^n u[n] + b^n u[n] \tag{E6.12.1}$$

La sequenza $c[n]$ è costituita dalla somma di due sequenze esponenziali *causali* e quindi la trasformata cercata è, come nell'esempio precedente,

$$C(z) = \frac{1}{1-bz^{-1}} + \frac{1}{1-az^{-1}} = \frac{2-(a+b)z^{-1}}{\left(1-az^{-1}\right)\left(1-bz^{-1}\right)} \tag{E6.12.2}$$

In questo caso la zona di convergenza *esiste certamente*. Infatti l'intersezione fra le regioni di convergenza delle trasformate $\mathcal{Z}$ delle due componenti *causali* di $c[n]$ infatti è pari a

$$|z| > \max(|a|, |b|) \tag{E6.12.3}$$

ed è rappresentata nella Figura 6.28 per a e b reali positivi, $b > a$.

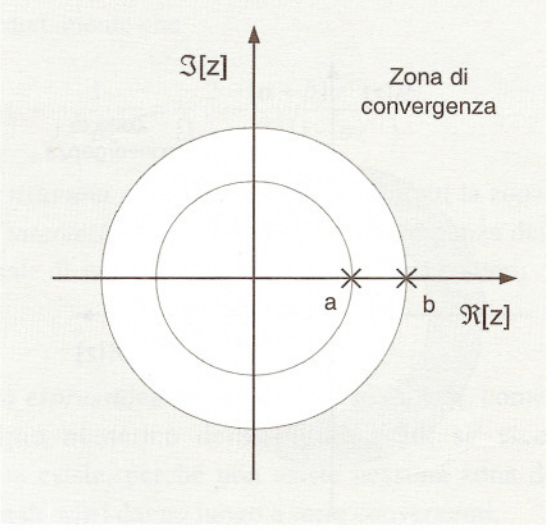


Figura 6.28 Zona di convergenza della trasformata nell'Esempio 6.12

6.3.2 Relazione con la trasformata di Fourier

Confrontiamo le espressioni della trasformata di Fourier $\overline{X}(f)$ e della trasformata $\mathcal{Z}$ $X(z)$ di una stessa sequenza $x[n]$:

$$X(z)=\sum_{n=-\infty}^{\infty}x[n]z^{-n} \quad , \quad \overline{X}(f)=\sum_{n=-\infty}^{\infty}x[n]e^{-j2\pi fT}=\sum_{n=-\infty}^{\infty}x[n]\left(e^{-j2\pi fT}\right)^{n} \tag{6.3.7}$$

È chiaro che, formalmente, si può ottenere la trasformata di Fourier dalla conoscenza della trasformata $\mathcal{Z}$ semplicemente ponendo

$$\overline{X}(f)=X(z)\big|_{z=e^{j2\pi fT}} \tag{6.3.8}$$

Questo significa valutare la $X(z)$ in tutti i punti del piano della variabile z che si trovano sulla *circonferenza di raggio unitario* $|z|=1$. Quando la frequenza f varia infatti tra $-1/2T$ e $1/2T$, la quantità $e^{j2\pi fT}$ "percorre" tale luogo geometrico a partire dal punto $-1+j0$, per ritornarvi dopo una intera rotazione in senso antiorario. Riguardo a questo punto, però, bisogna fare considerazioni simili a quelle viste nel Paragrafo 3.6 a proposito della relazione fra trasformata di Laplace e di Fourier a tempo continuo. Affinché la relazione $\overline{X}(f)=X(e^{j2\pi fT})$ abbia senso, ci si deve assicurare che la circonferenza $|z|=1$ *sia interamente contenuta nella zona di convergenza di* $X(z)$, come mostra la Figura 6.29, altrimenti la (6.3.8) non è applicabile.

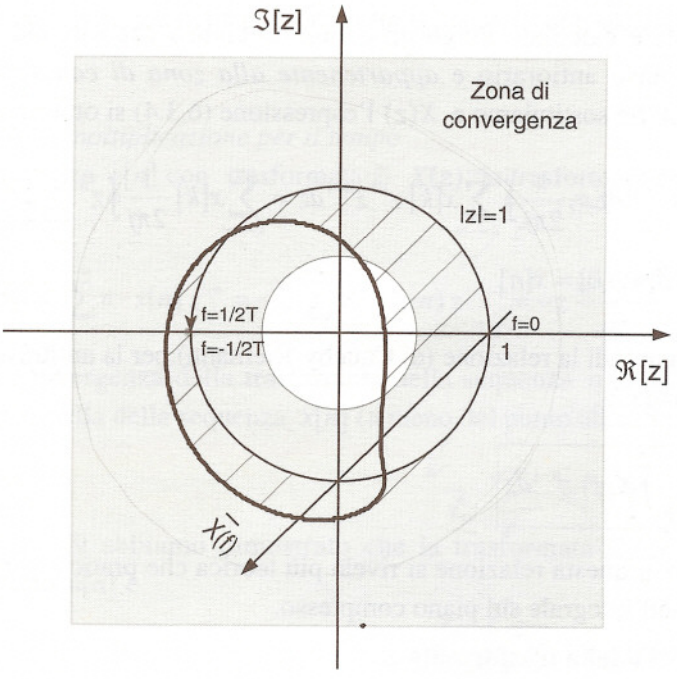


Figura 6.29 Estrazione, quando lecito, della trasformata di Fourier dalla trasformata $\mathcal{Z}$

6.3.3 Inversione della trasformata $\mathcal{Z}$

Qual è la relazione che permette di ricostruire una sequenza $x[n]$ a partire dall'espressione della sua trasformata $\mathcal{Z}$ $X(z)$? Il problema è equivalente a quello della ricostruzione di *coefficienti* della serie di Taylor-Laurent (6.3.4) nota la somma della serie stessa. Questi coefficienti possono essere ricavati attraverso il *teorema di Cauchy* applicato alla funzione di variabile complessa z^{n-1}. Se consideriamo un cammino di integrazione chiuso che circonda il punto $z = 0$ (ad esempio, la circonferenza $|z| = 1$) percorso in senso antiorario, il teorema di Cauchy stabilisce che

$$\frac{1}{2\pi j}\oint z^{n-1}dz = \delta[n] \tag{6.3.9}$$

Infatti, la funzione z^{n-1} ha un *polo* nel punto $z = 0$ per $n \leq 0$, ma l'unico caso in cui il *residuo* in questo polo è diverso da zero (e pari a 1) è quello in cui $n = 0$. Consideriamo allora la seguente espressione:

$$\frac{1}{2\pi j}\oint X(z)\, z^{n-1}dz \tag{6.3.10}$$

ove l'integrale è calcolato lungo un cammino chiuso che circonda l'origine, percorso in senso antiorario e *appartenente alla zona di convergenza* della funzione $X(z)$. Se sostituiamo a $X(z)$ l'espressione (6.3.4) si ottiene

$$\frac{1}{2\pi j}\oint X(z)\, z^{n-1}dz = \frac{1}{2\pi j}\oint \sum_{k=-\infty}^{+\infty} x[k]\, z^{-k} z^{n-1}dz = \sum_{k=-\infty}^{+\infty} x[k]\, \frac{1}{2\pi j}\oint z^{n-1-k}dz$$
$$= \sum_{k=-\infty}^{+\infty} x[k]\, \delta[n-k] = x[n] \tag{6.3.11}$$

Riassumiamo quindi la relazione (di Cauchy-Riemann) per la antitrasformazione della trasformata $\mathcal{Z}$:

$$\boxed{x[n] = \frac{1}{2\pi j}\oint X(z)\, z^{n-1}dz} \tag{6.3.12}$$

L'importanza di questa relazione si rivela più teorica che pratica, per la difficile valutazione dell'integrale sul piano complesso.

6.3.4 Proprietà della trasformata $\mathcal{Z}$

Dalla definizione (6.3.4) segue immediatamente che l'operazione di trasformata $\mathcal{Z}$ gode della proprietà di *linearità*. Enunciamo adesso e dimostriamo alcune

ulteriori proprietà salienti di cui gode tale trasformata.

• *Teorema del ritardo*

Data una sequenza $x[n]$ con trasformata $\mathcal{Z}$ $X(z)$, la trasformata della sequenza ritardata di n_0 passi è data da

$$\sum_{n=-\infty}^{+\infty} x[n-n_0]\, z^{-n} = \sum_{n=-\infty}^{+\infty} x[n]\, z^{-(n+n_0)} = z^{-n_0} \sum_{n=-\infty}^{+\infty} x[n]\, z^{-n} = z^{-n_0} X(z) \tag{6.3.13}$$

La zona di convergenza della sequenza $x[n]$ e della sequenza $x[n-n_0]$ coincidono a meno, eventualmente, dell'origine (se $n_0 > 0$) e del punto all'infinito (se $n_0 < 0$).

Esempio 6.13

La trasformata $\mathcal{Z}$ della sequenza $\delta[n]$ è data da $\Delta(z) = 1$ e la sua zona di convergenza è tutto il piano complesso. In virtù del teorema del ritardo, la trasformata $\mathcal{Z}$ della sequenza $\delta[n-1]$ è allora $z^{-1}\,\Delta(z) = z^{-1}$ e la sua zona di convergenza è costituita da tutto il piano complesso esclusa l'origine. Analogamente, la trasformata $\mathcal{Z}$ della sequenza $\delta[n+1]$ è data da $z\,\Delta(z) = z$ e la sua zona di convergenza è costituita da tutto il piano complesso escluso il punto all'infinito. ❑

• *Teorema della moltiplicazione per il tempo*

Data una sequenza $x[n]$ con trasformata $\mathcal{Z}$ $X(z)$, la trasformata della sequenza $n \cdot x[n]$ è data da

$$\mathcal{Z}[n \cdot x[n]] = \sum_{n=-\infty}^{+\infty} n \cdot x[n]\, z^{-n} = -z \sum_{n=-\infty}^{+\infty} x[n]\,(-n)\, z^{-n-1} = -z \frac{d\,X(z)}{dz} \tag{6.3.14}$$

La zona di convergenza della trasformata della sequenza $n \cdot x[n]$ coincide, in generale, con quella della sequenza $x[n]$ (a meno del punto all'infinito).

Esempio 6.14

Nell'Esempio 6.9 abbiamo dimostrato che la trasformata $\mathcal{Z}$ della sequenza gradino unitario $u[n]$ è

$$U(z) = \frac{1}{1-z^{-1}} = \frac{z}{z-1} \tag{E6.14.1}$$

con $|z| > 1$. Per il teorema della moltiplicazione per il tempo, la trasformata della

sequenza *rampa unitaria* $x[n] = n \cdot u[n]$ è data da

$$X(z) = -z \frac{d\,U(z)}{dz} = -z \frac{d}{dz}\left[\frac{z}{z-1}\right] = \frac{z}{(z-1)^2} = \frac{z^{-1}}{\left(1-z^{-1}\right)^2} \qquad \text{(E6.14.2)}$$

La zona di convergenza di $X(z)$ coincide con quella di $U(z)$, poiché in entrambi i casi è presente il solo polo $z_p = 1$. ❑

• *Teorema della convoluzione*

Date due sequenze $x[n]$ e $y[n]$ con trasformate $\mathcal{Z}$ rispettivamente $X(z)$ e $Y(z)$, la trasformata $\mathcal{Z}$ della loro somma di convoluzione $x[n] \otimes y[n]$ è

$$\mathcal{Z}[x[n] \otimes y[n]] = \sum_{n=-\infty}^{+\infty} \{x[n] \otimes y[n]\}\, z^{-n} = \sum_{n=-\infty}^{+\infty} \sum_{k=-\infty}^{\infty} x[k]\, y[n-k]\, z^{-n}$$

$$= \sum_{k=-\infty}^{+\infty} x[k] \sum_{n=-\infty}^{+\infty} y[n-k]\, z^{-n} \qquad (6.3.15)$$

Poiché, per il teorema del ritardo, si ha

$$\sum_{n=-\infty}^{+\infty} y[n-k]\, z^{-n} = z^{-k}\, Y(z) \qquad (6.3.16)$$

la (6.3.15) diviene

$$\sum_{n=-\infty}^{+\infty} \{x[n] \otimes y[n]\}\, z^{-n} = \sum_{k=-\infty}^{+\infty} x[k]\, z^{-k}\, Y(z) = X(z)\, Y(z) \qquad (6.3.17)$$

La trasformata $\mathcal{Z}$ della somma di convoluzione tra due sequenze è dunque pari al prodotto delle trasformate $\mathcal{Z}$ delle sequenze stesse.

Si osservi che la zona di convergenza della sequenza somma di convoluzione è data in generale dalla *intersezione* fra le regioni di convergenza di $X(z)$ e $Y(z)$, salvo casi particolari.

6.4 Sistemi a tempo discreto regolati da equazioni alle differenze

6.4.1 Un caso di studio

Introduciamo l'argomento di questo paragrafo attraverso lo studio di un caso particolare. Consideriamo dunque il circuito *R-C* di Figura 6.30. Come già dimostrato nel Capitolo 4, la risposta in frequenza di questo SLS a tempo *continuo* è espressa da

$$H(f)=\frac{Y(f)}{X(f)}=\frac{1}{1+j2\pi f\alpha}\ ,\quad \alpha=RC \tag{6.4.1}$$

da cui segue:

$$Y(f)+j2\pi f\alpha\, Y(f)=X(f) \tag{6.4.2}$$

Se si antitrasformano entrambi i membri della (6.4.2) si ricava la seguente equazione differenziale che caratterizza il comportamento del sistema nel dominio del tempo:

$$y'(t)=-\frac{1}{\alpha}y(t)+\frac{1}{\alpha}x(t)\quad \text{con}\quad y'(t)\triangleq dy(t)/dt \tag{6.4.3}$$

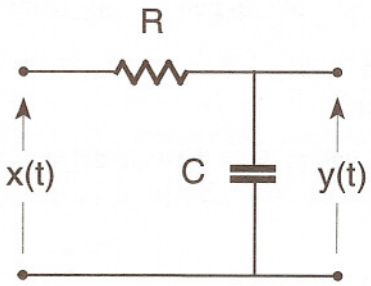


Figura 6.30 Circuito elettrico *R-C*

Le caratteristiche di *linearità* e *stazionarietà* del sistema sono riflesse nella struttura dell'equazione differenziale che regola il comportamento del sistema stesso: l'equazione è *lineare* e presenta *coefficienti costanti* (nel tempo). Il sistema in esame è solo un esemplare molto semplice della grande classe di sistemi a tempo continuo che sono per l'appunto regolati da *equazioni differenziali lineari a coefficienti costanti.* Qual è l'equivalente a tempo discreto di questa grande classe?

Per rispondere a questa domanda, riconsideriamo l'esempio particolare della squadra *R-C*, e cerchiamo di ricavare un sistema a tempo discreto che, ricevendo in ingresso la sequenza prodotta dal *campionamento* del segnale $x(t)$ con periodo T, fornisca in uscita una sequenza quanto più possibile simile a quella prodotta dal campionamento di $y(t)$ con lo stesso periodo. Risolviamo cioè un problema di *approssimazione* di un sistema a tempo continuo con uno a tempo discreto (chiamato problema di *simulazione*) schematizzato nella Figura 6.31.

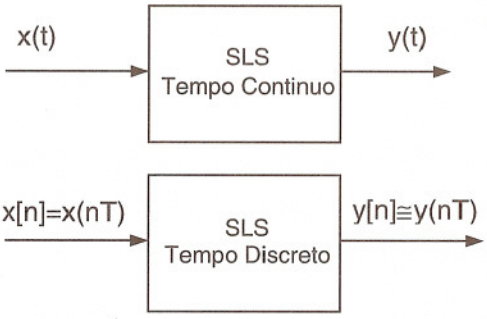


Figura 6.31 Simulazione di un SLS a tempo continuo con un SLS a tempo discreto

Cominciamo dunque con il riscrivere l'equazione differenziale (6.4.3) agli istanti nT relativi al campionamento dei segnali a tempo continuo:

$$y'(nT) = -\frac{1}{\alpha} y(nT) + \frac{1}{\alpha} x(nT) \tag{6.4.4}$$

Approssimiamo adesso il valore della derivata del segnale d'uscita all'istante nT con il valore del *rapporto incrementale* all'indietro calcolato su di un intervallo di campionamento:

$$y'(nT) \cong \frac{y(nT) - y((n-1)T)}{T} \tag{6.4.5}$$

Se usiamo questa approssimazione nella (6.4.4) e sostituiamo ai valori dei segnali a tempo continuo campionati i valori delle *sequenze* di ingresso/uscita del sistema a tempo discreto, otteniamo la seguente equazione:

$$\frac{y[n] - y[n-1]}{T} = -\frac{1}{\alpha} y[n] + \frac{1}{\alpha} x[n] \tag{6.4.6}$$

Questa relazione caratterizza dunque un sistema a tempo discreto il cui comportamento, nei limiti di validità dell'approssimazione numerica (6.4.5), è simile a, cioè *simula*, quello del sistema analogico di partenza.

La relazione (6.4.6) è una *equazione alle differenze*, cioè una relazione che lega i valori all'istante corrente n dell'ingresso e dell'uscita del sistema con le *differenze* tra tali valori e i rispettivi valori a istanti diversi ($n-1$, $n-2$ ecc.). Le equazioni alle differenze sono la controparte immediata a tempo discreto delle equazioni differenziali a tempo continuo. La relazione (6.4.6) appena ricavata, che possiamo riformulare come segue,

$$y[n]=\left(\frac{\alpha}{\alpha+T}\right)y[n-1]+\left(\frac{T}{T+\alpha}\right)x[n] \tag{6.4.7}$$

è caratterizzata dalle seguenti proprietà:

- è *lineare* poiché i valori delle sequenze $x[n]$ e $y[n]$ compaiono solo come termini di una *combinazione lineare* (su di essi cioè non viene svolta alcuna operazione *nonlineare*, come elevazione a potenza ecc.);
- è a *coefficienti costanti* poiché tutti i coefficienti della combinazione lineare suddetta *non* dipendono dalla variabile temporale n;
- è di *ordine 1* poiché la sequenza di uscita $y[n]$ compare ritardata al massimo di 1 passo: $y[n-1]$.

6.4.2 Implementazione con componenti elementari e generalizzazione

Possiamo *rappresentare* con un diagramma a blocchi l'equazione alle differenze (6.4.7), ottenendo, del sistema a tempo discreto, una descrizione che considereremo analoga allo "schema elettrico" *R-C* del sistema di partenza, nel senso che ugualmente ne dà una descrizione di tipo *grafico*. Per costruire questo diagramma a blocchi useremo i tre *componenti elementari* rappresentati in Figura 6.32, cioè l'*amplificatore ideale*, il *ritardatore* e il *sommatore*.

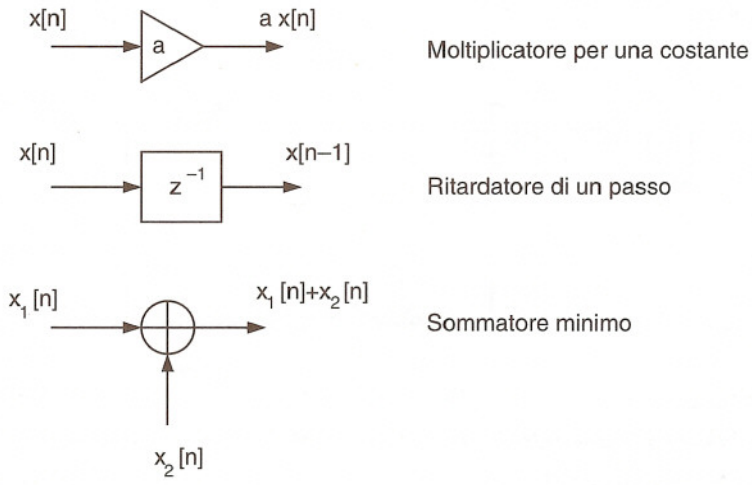


Figura 6.32 Componenti elementari a tempo discreto

Lo schema che descrive il comportamento del SLS a tempo discreto in esame è facilmente ricavabile dall'Equazione (6.4.7) ed è rappresentato in Figura 6.33. L'Equazione (6.4.7) è tipicamente *ricorsiva*: il valore corrente dell'uscita all'i-

stante n è ricavabile in funzione del valore dell'uscita al passo precedente $n-1$, che è ricavabile in funzione del valore dell'uscita al passo precedente $n-2$, che è ricavabile in funzione del valore dell'uscita al passo precedente $n-3$... Questa *ricorsione* si traduce dal punto di vista grafico nella "reazione di segnale" visibile in Figura 6.33, cioè nel percorso che "riporta" il segnale di uscita, opportunamente ritardato e scalato, verso l'ingresso.

Cerchiamo adesso di *estendere* i concetti visti per l'esempio particolare al caso generale di un SLS a tempo discreto che può essere descritto da una equazione alle differenze lineare e a coefficienti costanti di ordine N. Limitandoci a considerare soltanto sistemi *causali*, il comportamento del sistema è descritto da un'equazione del tipo

$$\sum_{m=0}^{N} a_m \, y[n-m] = \sum_{k=0}^{M} b_k \, x[n-k] \tag{6.4.8}$$

Senza perdere di generalità possiamo porre $a_0 = 1$ nella relazione e trovare la *forma normale* dell'equazione alle differenze:

$$y[n] = -\sum_{m=1}^{N} a_m \, y[n-m] + \sum_{k=0}^{M} b_k \, x[n-k] \tag{6.4.9}$$

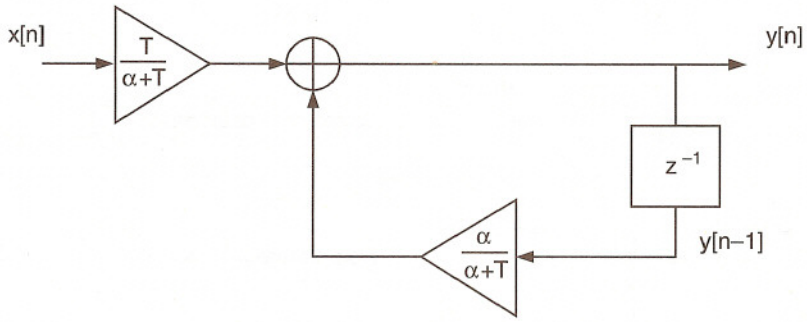


Figura 6.33 Schema a blocchi di un SLS del primo ordine

La causalità è adesso evidente: il valore assunto dalla sequenza di uscita all'istante n è determinato soltanto dai valori dell'ingresso e dell'uscita stessa per istanti *non successivi* a n. L'*ordine* dell'equazione coincide in pratica con il massimo ritardo con il quale compare la sequenza di uscita.

Utilizzando i componenti elementari di Figura 6.32, possiamo dare la rappresentazione generale di questa equazione mostrata in Figura 6.34, che prende il

nome di *rappresentazione in forma diretta*. Questa comporta la presenza di due *registri di ritardo* (o registri a scorrimento) di ordine M ed N; tali registri permettono di ottenere i valori ritardati rispettivamente dell'ingresso e dell'uscita necessari alla implementazione delle sommatorie nella (6.4.9), per un totale di $M+N$ elementi di ritardo. Il disegno fa riferimento al caso particolare $M > N$, ma in generale tra questi due parametri non c'è alcuna relazione.

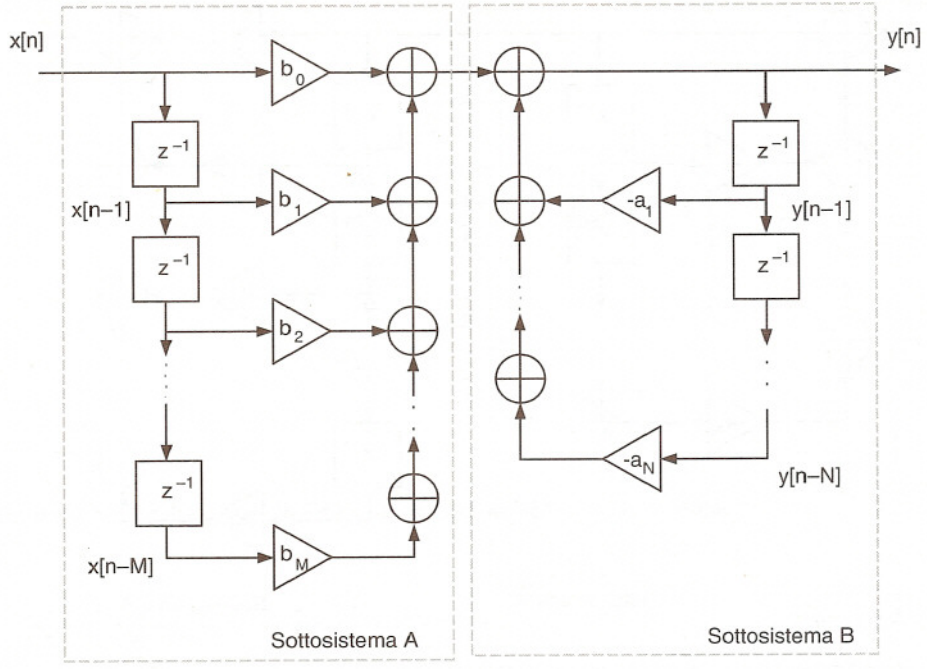


Figura 6.34 Realizzazione in forma diretta di un SLS causale di ordine N

Nella Figura 6.34 sono stati messi in evidenza due *sottosistemi* lineari e stazionari del SLS nella sua globalità, racchiusi nei riquadri a tratteggio. È allora possibile *invertire* l'ordine di tali sottosistemi senza che il comportamento globale muti minimamente: si ottiene così la struttura modificata rappresentata in Figura 6.35. Si nota però che in questa nuova configurazione i due registri di ritardo hanno in ingresso *il medesimo segnale*; essi possono essere sostituiti da un *unico* registro che "serve" entrambi i sottosistemi e di lunghezza pari al massimo tra M ed N. Si ottiene quindi la *realizzazione in forma canonica* del sistema mostrata in Figura 6.36. La struttura canonica minimizza il numero di *ritardi* necessari all'implementazione dell'equazione alle differenze (ovviamente si ha che, $\max(N,M) \leq N+M$), ma richiede la considerazione esplicita del segnale $w[n]$

interno al circuito (una sorta di "variabile di stato"). In un programma per calcolatore che implementa l'equazione alle differenze (6.4.9), la lunghezza totale dei registri di ritardo è pari al numero di *locazioni di memoria* che si devono riservare per i valori ritardati dei segnali di ingresso/uscita necessari per calcolare il valore corrente dell'uscita.

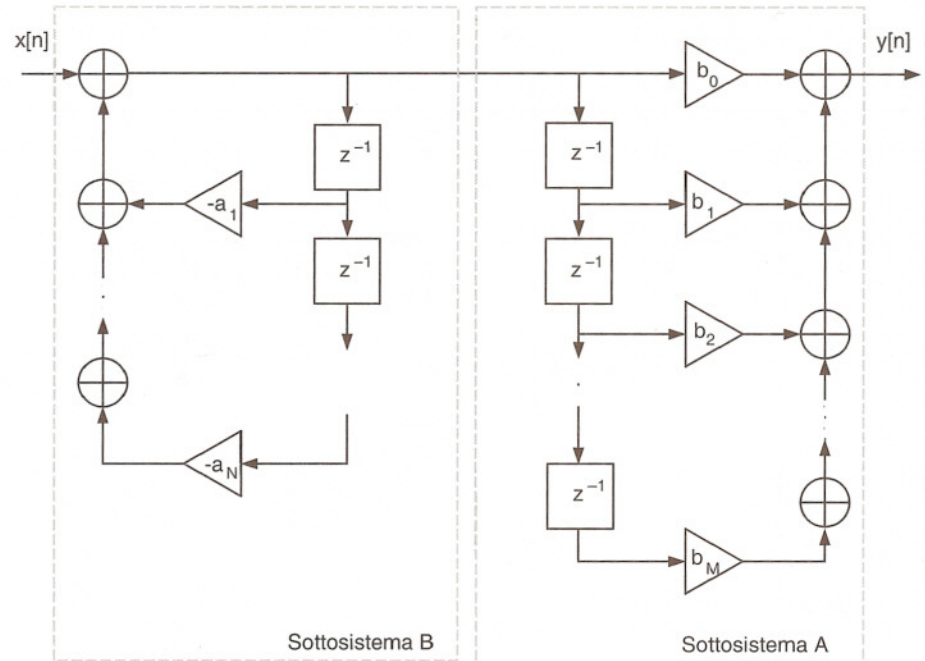


Figura 6.35 Modifica della forma diretta di Figura 6.34

6.4.3 Calcolo della risposta impulsiva

L'equazione alle differenze (6.4.9) o uno degli schemi a blocchi delle Figure 6.34-6.36 descrivono completamente il comportamento del SLS a tempo discreto nel dominio del tempo. Tuttavia, lo stesso sistema è completamente caratterizzato anche quando se ne conosce la risposta impulsiva

$$h[n] = \mathcal{T}[\delta[n]] \tag{6.4.10}$$

Il calcolo della sequenza $h[n]$ per il sistema dato può essere effettuato in maniera *ricorsiva* utilizzando l'equazione alle differenze in forma normale (6.4.9), con $x[n] = \delta[n]$ e $y[n] = h[n]$:

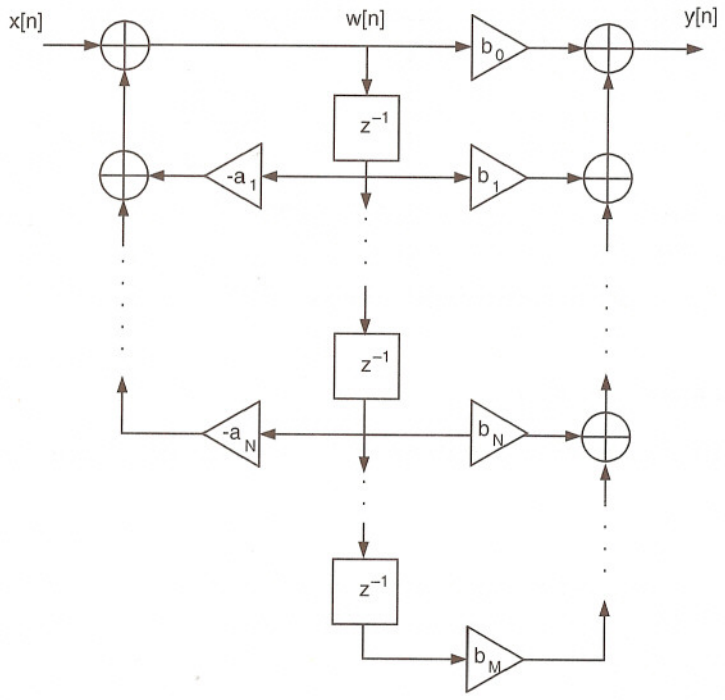


Figura 6.36 Realizzazione in forma canonica di un SLS causale di ordine N

$$h[n] = -\sum_{m=1}^{N} a_m \, h[n-m] + \sum_{k=0}^{M} b_k \, \delta[n-k] \tag{6.4.11}$$

Questa equazione può essere materialmente risolta con un calcolatore a partire dall'istante $n = 0$ e tenendo conto che, per la causalità del sistema, $h[n] = 0$ per $n < 0$.

Esempio 6.15

Consideriamo un sistema causale del primo ordine descritto dall'equazione alle differenze

$$y[n] = a\,y[n-1] + b\,x[n] \tag{E6.15.1}$$

la cui realizzazione in forma canonica è mostrata in Figura 6.37. Possiamo anche caratterizzare il sistema mediante la sua risposta impulsiva utilizzando la relazione ricorsiva che si ottiene ponendo $x[n] = \delta[n]$ e $y[n] = h[n]$ nella (E6.15.1):

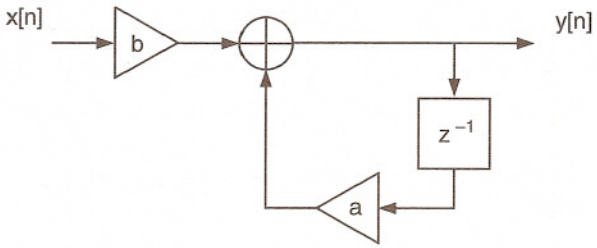


Figura 6.37 Forma canonica del sistema dell'Esempio 6.15

$$h[n] = a\, h[n-1] + b\, \delta[n] \tag{E6.15.2}$$

In particolare, ricordando che $h[n]$ è causale (e quindi $h[-1] = 0$), per $n = 0$ si ha

$$h[0] = b \tag{E6.15.3}$$

mentre per $n > 0$

$$h[n] = a\, h[n-1] \tag{E6.15.4}$$

La Tabella 6.1 riassume i calcoli che si devono effettuare per ricavare l'espressione della risposta impulsiva.

Tabella 6.1 Calcolo ricorsivo della risposta impulsiva

n	$\delta[n]$	$h[n] = a\,h[n-1] + b\,\delta[n]$
-1	0	0
0	1	b
1	0	$a \cdot b + 0 = ab$
2	0	$a \cdot ab + 0 = a^2 b$
3	0	$a \cdot a^2 b + 0 = a^3 b$
⋮	⋮	⋮
n	0	$a \cdot a^{n-1} b + 0 = a^n b$

Riassumendo i risultati della Tabella 6.1, concludiamo che

$$h[n] = b\, a^n u[n] \tag{E6.15.5}$$

❑

6.4.4 La funzione di trasferimento

La *funzione di trasferimento* di un SLS avente risposta impulsiva $h[n]$ è definita come la trasformata $\mathcal{Z}$ $H(z)$ della sequenza $h[n]$:

$$H(z) \triangleq \sum_{n=-\infty}^{+\infty} h[n]\, z^{-n} \tag{6.4.12}$$

Abbiamo dimostrato che, nota la risposta impulsiva di un sistema a tempo discreto, la sequenza di uscita $y[n]$ è data dalla somma di convoluzione tra la risposta impulsiva stessa e la sequenza di ingresso $x[n]$, cioè

$$y[n] = x[n] \otimes h[n] \tag{6.4.13}$$

Poiché la trasformata $\mathcal{Z}$ di una somma di convoluzione è data dal prodotto delle trasformate $\mathcal{Z}$ dei due segnali (vedi la (6.3.17)), dalla (6.4.13) si ha che

$$Y(z) = X(z)\, H(z) \tag{6.4.14}$$

quindi la funzione di trasferimento di un sistema è anche espressa dal *rapporto fra la trasformata $\mathcal{Z}$ della sequenza d'uscita e quella della sequenza d'ingresso*:

$$H(z) = \frac{Y(z)}{X(z)} \tag{6.4.15}$$

Torniamo ora a considerare i SLS causali descritti da equazioni alle differenze. L'equazione in forma normale è

$$y[n] = -\sum_{m=1}^{N} a_m\, y[n-m] + \sum_{k=0}^{M} b_k\, x[n-k] \tag{6.4.16}$$

Calcolando la trasformata $\mathcal{Z}$ di entrambi i membri ricaviamo:

$$Y(z) = -\sum_{m=1}^{N} a_m\, z^{-m} Y(z) + \sum_{k=0}^{M} b_k\, z^{-k} X(z) = -Y(z) \sum_{m=1}^{N} a_m\, z^{-m} + X(z) \sum_{k=0}^{M} b_k\, z^{-k} \tag{6.4.17}$$

da cui, con semplici passaggi,

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^{M} b_k z^{-k}}{1 + \sum_{m=1}^{N} a_m z^{-m}} \tag{6.4.18}$$

La funzione di trasferimento di un SLS causale descritto da una equazione alle differenze lineare a coefficienti costanti è quindi una *funzione razionale fratta* nella variabile z^{-1}.

Esempio 6.16
Consideriamo nuovamente il sistema causale descritto nell'Esempio 6.15:

$$y[n] = ay[n-1] + bx[n] \tag{E6.16.1}$$

Esso ricade nella forma generale (6.4.16) con $N=1$, $M=0$, $a_1 = -a$, $b_0 = b$. Si può allora ricavare direttamente la funzione di trasferimento del sistema

$$H(z) = \frac{b_0}{1 + a_0 z^{-1}} = \frac{b}{1 - az^{-1}} \tag{E6.16.2}$$

la cui antitrasformata, con zona di convergenza $|z| > |a|$, è

$$h[n] = b\, a^n u[n] \tag{E6.16.3}$$

Si osservi che la scelta della zona di convergenza deriva immediatamente dall'ipotesi di *causalità*. ❑

6.4.5 Sistemi a risposta impulsiva finita e infinita

Discutiamo ora alcuni casi particolari di SLS causali con funzione di trasferimento razionale fratta. Supponiamo, ad esempio, che sia $N=0$. L'equazione alle differenze (6.4.16) perde completamente il carattere di *ricorsività* e si semplifica nella equazione *non ricorsiva*

$$y[n] = \sum_{k=0}^{M} b_k\, x[n-k] \tag{6.4.19}$$

e la funzione di trasferimento corrispondente diviene un *polinomio* in z^{-1}:

$$H(z) = \sum_{k=0}^{M} b_k\, z^{-k} \tag{6.4.20}$$

Antitrasformando, si ricava immediatamente la risposta impulsiva del sistema:

$$h[n] = \sum_{k=0}^{M} b_k\, \delta[n-k] = \begin{cases} b_n & 0 \le n \le M \\ 0 & \textit{altrimenti} \end{cases} \tag{6.4.21}$$

I sistemi caratterizzati da una equazione alle differenze del tipo (6.4.19), e quindi non ricorsivi, sono chiamati sistemi FIR (*Finite Impulse Response*) in quanto, come mostra la (6.4.21), hanno una risposta impulsiva di *durata finita*. La forma diretta o canonica di un sistema FIR è rappresentata in Figura 6.38, ove si nota l'assenza di *reazione di segnale* dovuta alla mancata ricorsività dell'equazione alle differenze.

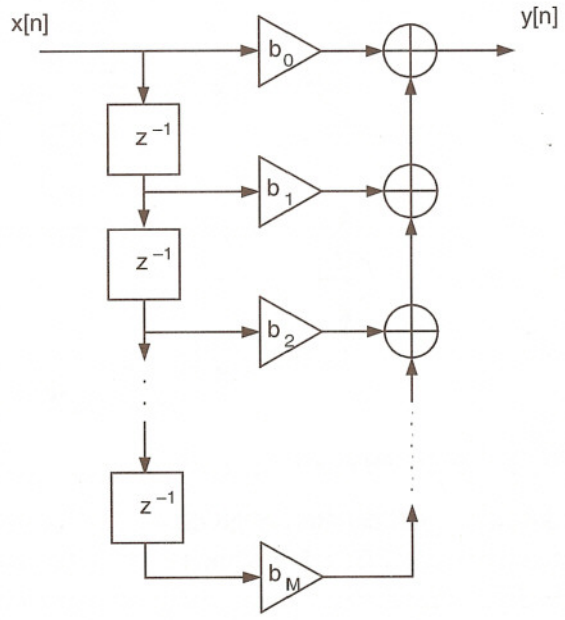


Figura 6.38 Realizzazione di un filtro FIR

Supponiamo ora che sia $N \neq 0$, come nel caso degli Esempi 6.15-6.16, in cui $N = 1$. L'equazione alle differenze comporta un calcolo ricorsivo: il valore della sequenza di uscita all'istante n è determinato dal valore all'istante precedente che è determinato dal valore all'istante precedente che è determinato dal valore all'istante precedente.... Questa ricorsione dà luogo a una *memoria infinita* per il sistema o, equivalentemente, a una *risposta impulsiva di durata infinita* (si vedano ancora gli Esempi 6.15-6.16). Per questa ragione i sistemi a tempo discreto caratterizzati da un'equazione alle differenze (6.4.16) con $N > 0$ sono detti IIR (*Infinite Impulse Response*) e la relativa funzione di trasferimento è una funzione razionale fratta nella variabile z^{-1}. Si introduce talvolta la nomenclatura di *puramente ricorsivo* per un sistema avente $M = 0$, cioè caratterizzato dall'equazione

$$y[n] = -\sum_{m=1}^{N} a_m \, y[n-m] + b_0 \, x[n] \tag{6.4.22}$$

corrispondente alla realizzazione di Figura 6.39.

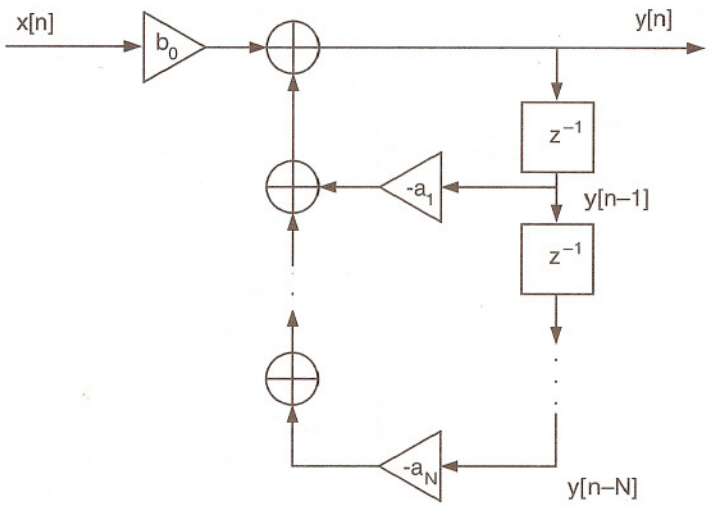


Figura 6.39 Realizzazione di un sistema puramente ricorsivo

Vogliamo ora evidenziare quali caratteristiche deve avere la funzione di trasferimento di un sistema causale FIR o IIR affinché sia garantita la stabilità secondo il criterio BIBO. Consideriamo innanzitutto un filtro FIR: poiché la risposta impulsiva $h[n]$ è costituita da un numero *finito* di valori di ampiezza limitata, la condizione (6.1.13) di assoluta sommabilità di $h[n]$ è sicuramente verificata. Segue che tutti i sistemi FIR sono sempre stabili senz'alcuna condizione sul numero dei coefficienti e sui valori (limitati) da essi assunti: per questa ragione si usa dire che i sistemi (filtri) FIR sono *incondizionatamente stabili.*

Per i filtri IIR, invece, la stabilità non è sempre garantita, come si vede facilmente con un semplice esempio.

Esempio 6.17

Consideriamo lo stesso sistema dell'Esempio 6.15, con $a = 2$ e $b = 1$:

$$y[n] = 2y[n-1] + x[n] \tag{E6.17.1}$$

e ricalcoliamone la risposta impulsiva. Il sistema è chiaramente IIR perché

ricorsivo, e la risposta impulsiva si calcola attraverso la Tabella 6.2.

Tabella 6.2 Calcolo ricorsivo della risposta impulsiva

n	$\delta[n]$	$h[n]=2h[n-1]+\delta[n]$
-1	0	0
0	1	1
1	0	$2+0=2$
2	0	$2\cdot 2+0=4$
3	0	$2\cdot 4+0=8$
⋮	⋮	⋮
n	0	$2\cdot 2^{n-1}+0=2^n$

Riassumendo i risultati della Tabella 6.2, concludiamo che

$$h[n]=2^n u[n] \tag{E6.17.2}$$

che diverge quando $n\to\infty$ e non è assolutamente sommabile. Il semplice sistema IIR dunque *non* è stabile. ❑

Dobbiamo trovare un criterio che permetta di stabilire se un dato sistema IIR è stabile o meno. Analizziamo quindi il caso generale di un sistema IIR puramente ricorsivo con funzione di trasferimento

$$H(z)=\frac{b_0}{1+a_1z^{-1}+a_2z^{-2}+\ldots+a_Nz^{-N}} \tag{6.4.23}$$

e calcoliamo per prima cosa le N radici dell'equazione in z:

$$1+a_1z^{-1}+a_2z^{-2}+\ldots+a_Nz^{-N}=0 \tag{6.4.24}$$

cioè i *poli* z_{p_m}, $m=1,\ldots,N$ della funzione di trasferimento. Supponendo per semplicità che i poli siano tutti a molteplicità unitaria, possiamo poi riscrivere la (6.4.23) attraverso la *scomposizione in fratti semplici*:

$$H(z)=\sum_{m=1}^{N}\frac{A_m}{1-z_{p_m}z^{-1}} \quad , \quad A_m=H(z)\left(1-z_{p_m}z^{-1}\right)\Big|_{z=z_{p_m}} \tag{6.4.25}$$

Antitrasformando la (6.4.25) si ricava la seguente espressione per la risposta impulsiva del sistema:

$$h[n] = \sum_{m=1}^{N} h_m[n] \tag{6.4.26}$$

ove

$$h_m[n] = A_m \, z_{p_m}^n \, u[n] \tag{6.4.27}$$

Come sappiamo, condizione necessaria e sufficiente per la stabilità BIBO è l'assoluta sommabilità della risposta impulsiva. In questo caso tale condizione è verificata se e solo se

$$\sum_{n=-\infty}^{\infty} |h_m[n]| < +\infty, \quad m = 1, \ldots, N \tag{6.4.28}$$

cioè se e solo se

$$\sum_{n=0}^{+\infty} |z_{p_m}|^n < +\infty \ , \quad m = 1, \ldots, N \tag{6.4.29}$$

Il sistema è allora stabile sotto la *condizione*

$$|z_{p_m}| < 1 \ , \quad m = 1, \ldots, N \tag{6.4.30}$$

ovvero sotto la condizione che *i poli della funzione di trasferimento siano tutti interni alla circonferenza di raggio unitario* $|z| = 1$ del piano z.

Per un sistema causale e *stabile*, dunque, la zona di convergenza è la parte di piano esterna alla circonferenza $|z| = |z_{p,Max}|$, ove $z_{p,Max}$ è quello tra gli N poli avente modulo massimo. Poiché per la stabilità deve essere $|z_{p,Max}| < 1$, necessariamente la circonferenza di raggio unitario $|z| = 1$ è *interamente contenuta nella zona di convergenza* della funzione di trasferimento. Dalle considerazioni sulla relazione fra trasformate di Fourier e $\mathcal{Z}$ del Paragrafo 6.3, segue che un sistema causale e *stabile* ha risposta in frequenza pari a

$$\overline{H}(f) = H(z)\big|_{z=e^{j2\pi fT}} \tag{6.4.31}$$

6.5 Cenni al progetto di filtri numerici IIR

Una delle funzioni elementari di elaborazione dei segnali a tempo discreto che è frequentemente necessaria in una varietà di problemi pratici è il *filtraggio*. I tipi-base di filtri a tempo discreto sono già stati rapidamente menzionati nel Paragrafo 6.3, ma, come nel caso del tempo continuo, conducono a sistemi non

fisicamente realizzabili. Sorge quindi la questione di realizzare filtri a tempo discreto che rispettino specifiche di selettività, ma che risultino di facile progettazione e implementazione.

Un criterio di progetto semplice e che conduce a risultati in molti casi soddisfacenti è quello di realizzare filtri a tempo discreto partendo da *prototipi* di filtri analogici, cioè a tempo continuo. In un certo senso, questo è il punto di vista adottato nel caso di studio del Paragrafo 6.4.1 per ricavare l'equazione alle differenze di un SLS a tempo discreto, partendo da un SLS a tempo continuo. Le tecniche che esporremo sono diverse da quella dell'esempio ma sono accomunate da una proprietà importante. Se esse vengono applicate a filtri analogici causali la cui funzione di trasferimento $H_a(s)$ (intesa come trasformata di Laplace della risposta impulsiva $h_a(t)$) è una funzione razionale fratta in s, danno luogo a filtri causali a tempo discreto la cui funzione di trasferimento è razionale fratta in z^{-1}, e quindi implementabile con una delle strutture viste nel Paragrafo 6.4. Poiché le caratteristiche dei filtri analogici sono ben note, il progetto di un filtro a tempo discreto di caratteristiche date diventa quindi un compito relativamente semplice.

6.5.1 La tecnica dell'invarianza impulsiva

Il punto di partenza per la tecnica dell'*invarianza impulsiva* è rappresentato dalla descrizione di un sistema a tempo continuo attraverso la risposta impulsiva $h_a(t)$. Noto questo segnale, si ottiene la risposta impulsiva $h[n]$ del SLS *a tempo discreto* attraverso il *campionamento della risposta a tempo continuo* $h_a(t)$ con un opportuno periodo T. In altre parole si impone che

$$h[n] = T\, h_a(nT) \tag{6.5.1}$$

ove la moltiplicazione per il fattore T viene introdotta, oltre che per motivi di carattere "dimensionale", per far sì che le risposte in frequenza dei due sistemi siano il più possibile simili. Infatti la risposta in frequenza del filtro a tempo discreto $\overline{H}(f)$ è legata alla risposta in frequenza del filtro analogico $H_a(f)$ attraverso la relazione (vedi la (5.4.7))

$$\overline{H}(f) = \sum_{k=-\infty}^{+\infty} H_a\left(f - \frac{k}{T}\right) \tag{6.5.2}$$

Se il segnale $h_a(t)$ è rigorosamente limitato in banda e il periodo di campionamento T rispetta la condizione di Nyquist, allora il filtro a tempo discreto ha le stesse caratteristiche di selettività del filtro analogico poiché il campionamento

non dà luogo ad aliasing. Nel periodo "base" $[-1/2T, 1/2T]$, quindi, la risposta in frequenza del filtro a tempo discreto ha esattamente lo stesso andamento di quella del filtro analogico "prototipo", come illustra la Figura 6.40 per un filtro passa-basso ideale.

Per evitare problemi di realizzabilità, i prototipi analogici che si considerano sono filtri *causali* e la loro risposta in frequenza non può essere rigorosamente limitata in banda. Un certo ammontare di aliasing sarà comunque presente, come è esemplificato dalla Figura 6.41 che rappresenta la risposta in frequenza (per semplicità, a valori reali) di un filtro passa-basso analogico reale e quella del corrispondente filtro a tempo discreto. Per questa ragione, l'andamento di $\overline{H}(f)$, nell'intervallo frequenziale di interesse $[-1/2T, 1/2T]$, risulta diverso da quello della funzione $H_a(f)$ nel medesimo intervallo. Tuttavia, se la frequenza di campionamento $1/T$ è sufficientemente grande in rapporto alla "banda" del segnale $h_a(t)$, allora l'effetto dell'aliasing può essere trascurabile, e la funzione $\overline{H}(f)$ può risultare molto simile, nella banda di interesse, alla funzione $H_a(f)$.

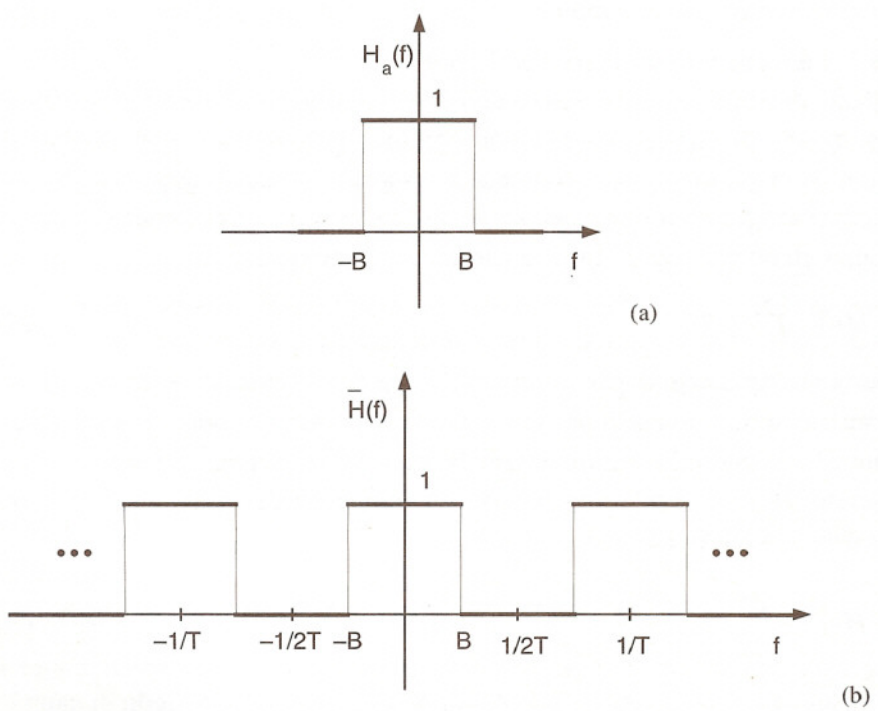


Figura 6.40 Risposte in frequenza di un filtro analogico passa-basso ideale (a), e del filtro a tempo discreto ricavato con l'invarianza impulsiva e con $f_c = 1/T = 4B$ (b)

È importante osservare che la relazione (6.5.2) limita l'applicabilità della tecnica dell'invarianza impulsiva alle classi dei filtri la cui risposta in frequenza ha un *limite superiore di banda*, anche inteso in senso approssimato (banda a –3 dB, banda efficace). Non ha quindi senso applicare l'invarianza impulsiva a filtri passa-alto o elimina-banda per i quali l'aliasing è intrinsecamente *ineliminabile*.

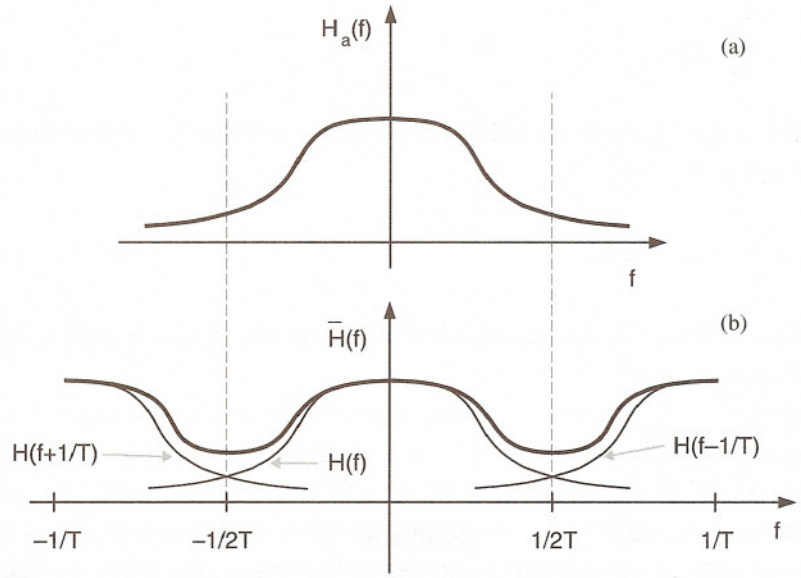


Figura 6.41 Risposte in frequenza di un filtro analogico passa-basso reale (a), e del filtro a tempo discreto ricavato con l'invarianza impulsiva (b)

Se il filtro prototipo analogico ha una funzione di trasferimento $H_a(s)$ razionale fratta in s, si può ricavare direttamente la funzione di trasferimento $H(z)$ del filtro a tempo discreto (che risulta a sua volta razionale fratta in z^{-1}), senza dover esplicitamente calcolare la risposta impulsiva a tempo continuo $h_a(t)$. Consideriamo infatti la generica funzione di trasferimento $H_a(s)$ di un sistema analogico descritto da una equazione differenziale lineare a coefficienti costanti di ordine N:

$$H_a(s) = \frac{N(s)}{D(s)} = \frac{\beta_0 + \beta_1 s + \ldots + \beta_M s^M}{\alpha_0 + \alpha_1 s + \ldots + \alpha_N s^N} \tag{6.5.3}$$

con $M < N$ per semplicità. Una volta calcolati i *poli* s_i, $i = 1,\ldots,N$ di $H_a(s)$ risolvendo l'equazione $D(s) = 0$, la stessa funzione può essere scomposta in fratti semplici:

$$H_a(s) = \sum_{i=1}^{N} \frac{A_i}{s - s_i} \tag{6.5.4}$$

ove A_i è il *residuo* del polo s_i, $A_i = H_a(s)(s - s_i)\big|_{s=s_i}$ (i poli sono a molteplicità unitaria per semplicità). Antitrasformando questa espressione si ottiene la risposta impulsiva del sistema analogico:

$$h_a(t) = \sum_{i=1}^{N} A_i \, e^{s_i t} \, \mathrm{u}(t) \tag{6.5.5}$$

e campionando la risposta analogica si ricava la seguente espressione della risposta impulsiva $h[n]$:

$$h[n] = \sum_{i=1}^{N} T A_i e^{n s_i T} u[n] \tag{6.5.6}$$

Tramite trasformata $\mathcal{Z}$, si determina infine la funzione di trasferimento $H(z)$ del filtro a tempo discreto:

$$H(z) = \sum_{i=1}^{N} \frac{T A_i}{1 - e^{s_i T} z^{-1}} \tag{6.5.7}$$

In conclusione, la funzione di trasferimento $H(z)$ del filtro a tempo discreto è direttamente ricavabile dai poli e dai residui della funzione di trasferimento del filtro analogico. Si nota in particolare che i poli s_i del filtro analogico si trasformano nei poli $z_i = e^{s_i T}$ del sistema a tempo discreto. Se il sistema di partenza è stabile, esso ha poli a parte reale negativa e quindi i poli del sistema a tempo discreto così ottenuto hanno *modulo minore di uno*. Anche il filtro a tempo discreto progettato con la tecnica dell'invarianza impulsiva è quindi stabile.

Esempio 6.18

Supponiamo di voler progettare un filtro passa-basso a tempo discreto utilizzando il prototipo analogico di Figura 6.42.

La funzione di trasferimento del filtro analogico è

$$H_a(s) = \frac{1}{s\alpha + 1} = \frac{1/\alpha}{s + 1/\alpha} \tag{E6.18.1}$$

ove $\alpha = RC$ rappresenta la costante di tempo del circuito. È facile verificare che la funzione $H_a(s)$ è del tipo (6.5.4) con $N = 1$, $s_1 = -1/\alpha$ e $A_1 = 1/\alpha$. Allora la

funzione di trasferimento $H(z)$ del filtro numerico si ricava immediatamente dalla (6.5.7):

$$H(z)=\left(\frac{T}{\alpha}\right)\frac{1}{1-e^{-T/\alpha}z^{-1}} \qquad \text{(E6.18.2)}$$

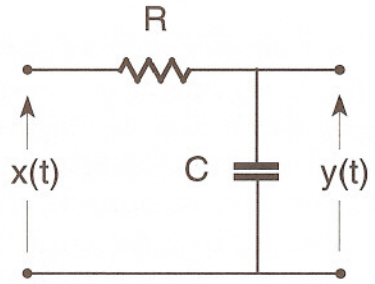


Figura 6.42 Filtro *R-C* prototipo

La struttura diretta del filtro è rappresentata in Figura 6.43. È importante osservare che, poiché il prototipo analogico non ha banda rigorosamente limitata, per ridurre l'effetto dell'aliasing è necessario scegliere opportunamente il valore della frequenza di campionamento $1/T$ nei confronti della banda del filtro $B_{-3}=1/(2\pi\alpha)$. Per questo filtro del primo ordine (avente bassa pendenza di attenuazione), è buona norma usare un valore di f_c *incrementato* di un fattore 5÷10 rispetto a quello previsto dalla condizione di Nyquist relativamente alla banda B_{-3} (cioè $f_c=10\div 20B_{-3}$)

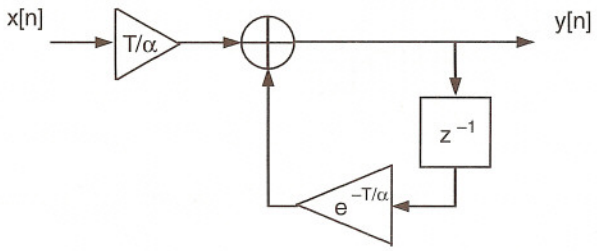


Figura 6.43 Filtro a tempo discreto ottenuto con il metodo dell'invarianza impulsiva

❑

Facciamo adesso delle ulteriori considerazioni sulla tecnica dell'invarianza impulsiva con riferimento all'esempio appena esposto. In particolare appli-

chiamo di nuovo a tale esempio la stessa tecnica di progetto, ma seguendo alla lettera il procedimento che prevede il campionamento della risposta impulsiva nel tempo. Quest'ultima si ricava facilmente dalla funzione di trasferimento (E6.18.1):

$$h_a(t) = \frac{1}{\alpha} e^{-t/\alpha}\, \mathrm{u}(t) \tag{6.5.8}$$

Ricaviamo poi la risposta impulsiva $h[n]$ del sistema a tempo discreto applicando la relazione (6.5.1) che richiede di campionare $h_a(t)$ per $t = nT$, come rappresentato in Figura 6.44. Si nota che l'operazione di campionamento comporta una *ambiguità* per $t = 0$ poiché, in questo istante, la funzione $h_a(t)$ è *discontinua*. Resta quindi incerto il valore da attribuire ad $h[0]$. Nel ricavare la relazione (6.5.6) abbiamo infatti scelto arbitrariamente il valore $h[0] = T h_a(0^+)$, cioè il limite *destro* del segnale. Nell'analisi di Fourier, il valore più opportuno da attribuire a un segnale discontinuo è la *semisomma* dei limiti destro e sinistro nel punto di discontinuità.

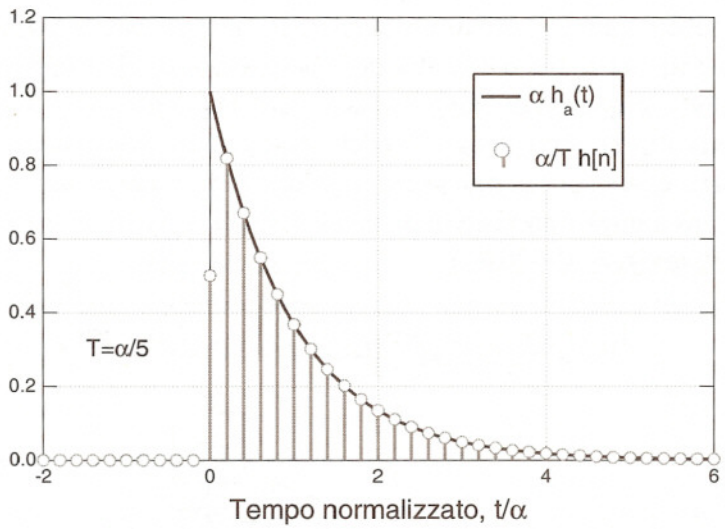


Figura 6.44 Campionamento di una risposta impulsiva con discontinuità

Se riesaminiamo il caso generale alla luce della precedente osservazione, dalla (6.5.5) si ha

$$h_a(0^-) = 0 \tag{6.5.9}$$

$$h_a(0^+) = \sum_{i=1}^{N} A_i \tag{6.5.10}$$

per cui si deve scegliere

$$h[0] = \frac{T}{2}\left(h_a(0^-) + h_a(0^+)\right) = \frac{T}{2} h_a(0^+) = \frac{T}{2}\sum_{i=1}^{N} A_i \tag{6.5.11}$$

come in Figura 6.44, e si deve correggere il valore in $n = 0$ della sequenza $h[n]$ (6.5.6):

$$\begin{aligned} h[n] &= \sum_{i=1}^{N} TA_i\, e^{ns_iT} u[n] - \frac{T}{2} h_a(0^+)\delta[n] \\ &= \sum_{i=1}^{N} T\, A_i\, e^{ns_iT} u[n] - \frac{T}{2}\sum_{i=1}^{N} A_i\, \delta[n] \end{aligned} \tag{6.5.12}$$

Quindi, la funzione di trasferimento più appropriata è

$$H(z) = \sum_{i=1}^{N} \frac{TA_i}{1 - e^{s_iT} z^{-1}} - \frac{T}{2}\sum_{i=1}^{N} A_i \tag{6.5.13}$$

che differisce dalla (6.5.7) per un termine *correttivo* che tiene conto dell'eventuale punto di discontinuità nell'origine.

Esempio 6.19

Riconsideriamo il progetto dell'Esempio 6.18 e riprogettiamo il filtro a tempo discreto tenendo conto della discontinuità nell'origine della risposta impulsiva del filtro analogico. Essendo, come già visto nell'esempio precedente, $N = 1$, $s_1 = -1/\alpha$ e $A_1 = 1/\alpha$, le relazioni (6.5.12)-(6.5.13) permettono di scrivere immediatamente le espressioni della risposta impulsiva e della funzione di trasferimento del filtro a tempo discreto:

$$h[n] = \frac{T}{\alpha} e^{-\frac{nT}{\alpha}} u[n] - \frac{T}{2\alpha}\delta[n] \tag{E6.19.1}$$

$$H(z) = \frac{T/\alpha}{1 - e^{-T/\alpha} z^{-1}} - \frac{T}{2\alpha} = \frac{T}{2\alpha}\left[\frac{2}{1 - e^{-T/\alpha} z^{-1}} - 1\right] = \frac{T}{2\alpha}\frac{1 + e^{-T/\alpha} z^{-1}}{1 - e^{-T/\alpha} z^{-1}} \tag{E6.19.3}$$

La nuova struttura del filtro in forma canonica è rappresentata in Figura 6.45. Si noti che il fattore $T/2\alpha$ comune ai coefficienti b_0 e b_1 è stato messo in evidenza

davanti al circuito. Per rimarcare il fenomeno dell'aliasing, riportiamo inoltre in Figura 6.46 l'andamento delle risposte in ampiezza e fase del filtro a tempo discreto (linee a tratto continuo)

$$\overline{H}(f) = H\left(e^{j2\pi fT}\right) = \frac{T}{2\alpha}\frac{1+e^{-\left(\frac{T}{\alpha}+j2\pi fT\right)}}{1-e^{-\left(\frac{T}{\alpha}+j2\pi fT\right)}} \qquad \text{(E6.19.4)}$$

confrontate con quelle del filtro analogico prototipo (linee punteggiate). ❑

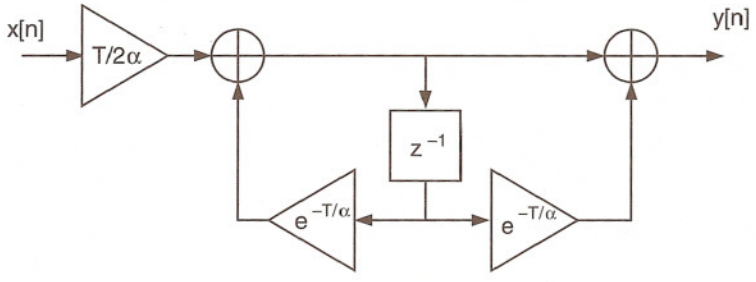


Figura 6.45 Versione corretta del filtro di Figura 6.43

6.5.2 La tecnica della trasformazione bilineare

L'esempio appena visto evidenzia i limiti di applicabilità della tecnica dell'invarianza impulsiva, che in particolare non è adatta quando è necessario rispettare alcune specifiche. Se ad esempio si vuole realizzare un filtro numerico di banda B data, ma non si può aumentare la frequenza di campionamento, l'effetto dell'aliasing porta una notevole differenza tra la risposta del prototipo e quella del filtro discreto, e la specifica non viene rispettata: si deve quindi cambiare il criterio di progetto del filtro a tempo discreto.

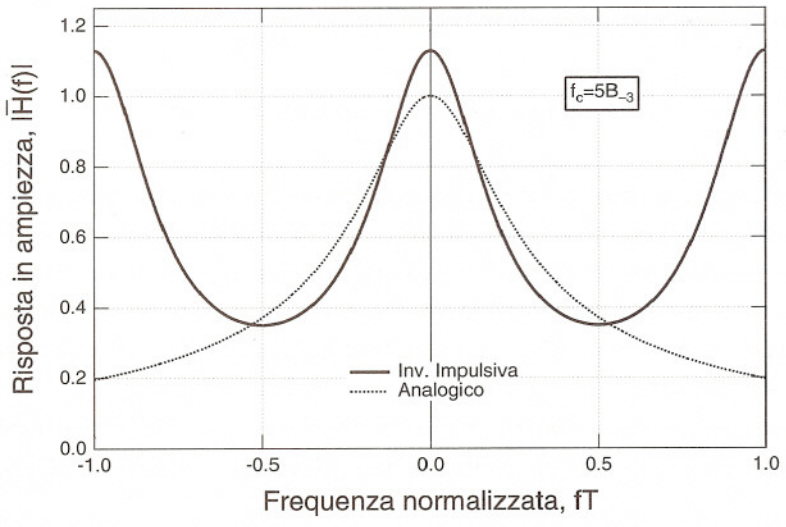
Supponiamo dunque di voler approssimare con un filtro a tempo discreto il comportamento del filtro R-C dell'Esempio 6.18, la cui funzione di trasferimento è ($\tau = RC$)

$$H_a(s) = \frac{1}{1+s\tau} \qquad (6.5.14)$$

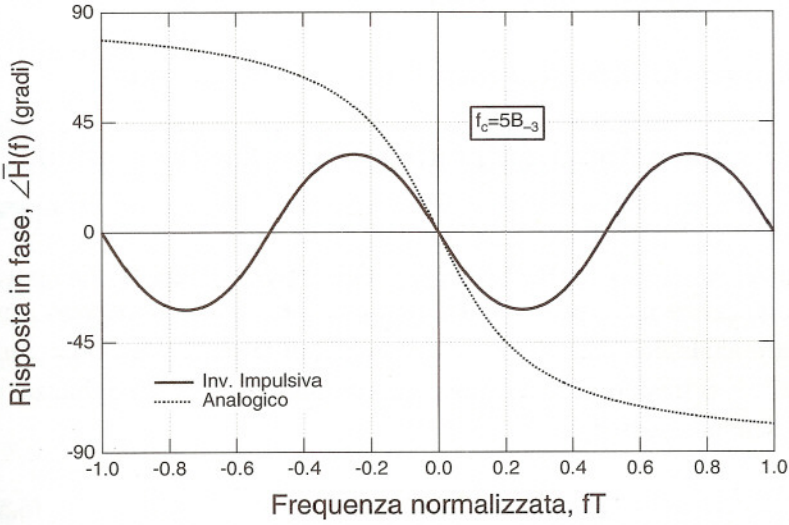
Da questa funzione di trasferimento è semplice ricavare l'*equazione differenziale* che regola il comportamento del filtro stesso:

$$Y(s)(1+s\tau) = X(s) \;\Rightarrow\; y(t) + \tau\frac{dy(t)}{dt} = x(t) \qquad (6.5.15)$$

CENTRO BIBLIOTECARIO INGEGNERIA PISA



(a)



(b)

Figura 6.46 Risposta in ampiezza (a) e in fase (b) del filtro di Figura 6.45

Integrando ambo i membri tra un instante iniziale t_0 e il generico istante t, l'equazione differenziale si trasforma in una equivalente *equazione integrale*:

$$\int_{t_0}^{t} y(\alpha)\, d\alpha + \tau\left[y(t) - y(t_0)\right] = \int_{t_0}^{t} x(\alpha)\, d\alpha \tag{6.5.16}$$

Valutiamo adesso questa relazione per $t_0 = (n-1)T$ e $t = nT$, ove T rappresenta l'intervallo di campionamento; la (6.5.16) diventa allora

$$\int_{(n-1)T}^{T} y(\alpha)\, d\alpha + \tau\big[y(nT) - y((n-1)T)\big] = \int_{(n-1)T}^{T} x(\alpha)\, d\alpha \tag{6.5.17}$$

A questo punto il nostro scopo è quello di passare da questa relazione tra i segnali *analogici* in ingresso/uscita al filtro *R-C,* a una relazione "equivalente" che leghi i segnali *discreti* $x[n]$ e $y[n]$ in ingresso/uscita a un filtro numerico il quale simuli il comportamento del filtro analogico. Cerchiamo allora di *approssimare* i due integrali che compaiono nella (6.5.17) con una formula numerica. Ad esempio, l'integrale a secondo membro può essere approssimato con la *regola del trapezio*: come è suggerito dalla Figura 6.47, l'area sottesa dal grafico della curva di $x(\alpha)$ è all'incirca uguale all'area del *trapezio* ABCD, cioè

$$\int_{(n-1)T}^{nT} x(\alpha)\, d\alpha \cong \frac{T}{2}\big[x(nT) + x((n-1)T)\big] \tag{6.5.18}$$

Applicando questa formula ai due integrali nella (6.5.17) si ottiene

$$\frac{T}{2}\big[y(nT) + y((n-1)T)\big] + \tau\big[y(nT) - y((n-1)T)\big] \cong \frac{T}{2}\big[x(nT) + x((n-1)T)\big] \tag{6.5.19}$$

Questa equazione, valida quando la regola del trapezio è una buona approssimazione, può essere usata per caratterizzare il sistema a tempo discreto equivalente al filtro analogico dato. Si introducono quindi le sequenze $y[n] = y(nT)$ e $x[n] = x(nT)$ e si trasforma la relazione approssimata in una uguaglianza, ottenendo con semplici passaggi:

$$y[n] = \frac{\tau - T/2}{\tau + T/2} y[n-1] + \frac{T/2}{\tau + T/2}\big[x[n] + x[n-1]\big] \tag{6.5.20}$$

che rappresenta l'*equazione alle differenze* del sistema a tempo discreto cercato. Applicando la trasformata $\mathcal{Z}$ all'Equazione (6.5.20), si può ricavare la corrispondente funzione di trasferimento del sistema:

$$H(z) = \frac{\dfrac{T/2}{\tau + T/2}\left(1 + z^{-1}\right)}{1 - \dfrac{\tau - T/2}{\tau + T/2} z^{-1}} = \frac{1}{1 + \dfrac{2}{T} \cdot \dfrac{1 - z^{-1}}{1 + z^{-1}}\, \tau} \tag{6.5.21}$$

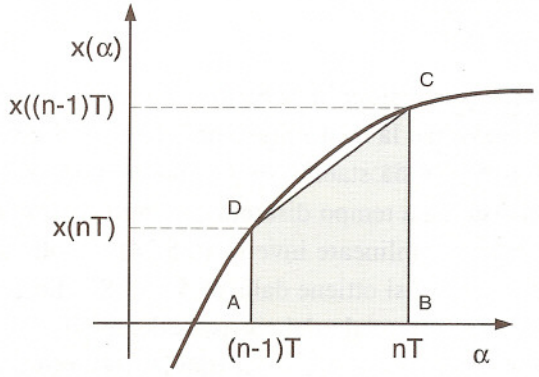


Figura 6.47 Integrazione numerica: regola del trapezio

che, confrontata con la propria controparte analogica $H(s)$ (6.5.14), porta alla conclusione che le due funzioni di trasferimento sono legate dalla relazione

$$H(z) = H_a\left(\frac{2}{T}\cdot\frac{1-z^{-1}}{1+z^{-1}}\right) \tag{6.5.22}$$

La relazione, ricavata per il caso particolare del filtro *R-C*, può essere poi applicata in generale per ricavare la funzione di trasferimento $H(z)$ del filtro a tempo discreto dalla funzione di trasferimento $H_a(s)$ di un qualunque prototipo analogico. Questa "sostituzione" di variabile ha il significato che emerge dall'esempio: ogniqualvolta nel sistema analogico si deve compiere un'operazione di integrazione, essa viene approssimata con la regola del trapezio per ricavare il filtro "simulatore" a tempo discreto. La relazione

$$s = \frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}} \tag{6.5.23}$$

dal punto di vista matematico istituisce una corrispondenza (trasformazione) fra i piani complessi s e z ed è nota come *trasformazione bilineare.*

Cerchiamo ora di stabilire le caratteristiche più importanti di questa trasformazione. Osserviamo innanzitutto che dalla (6.5.23) segue facilmente la relazione di corrispondenza inversa

$$z = \frac{1+sT/2}{1-sT/2} \tag{6.5.24}$$

che è a sua volta bilineare. Se poniamo $s = \sigma + j\omega$, si ottiene da quest'ultima

$$|z| = \sqrt{\frac{(1+\sigma T/2)^2 + (\omega T/2)^2}{(1-\sigma T/2)^2 + (\omega T/2)^2}} \tag{6.5.25}$$

che permette di trarre conclusioni importanti sulla *stabilità* del sistema a tempo discreto realizzato attraverso la trasformazione bilineare. Supponiamo infatti che il sistema analogico $H_a(s)$ sia stabile, quindi possieda poli s_{p_i} con parte reale negativa. I poli del sistema a tempo discreto z_{p_i} sono chiaramente ottenibili applicando la trasformazione bilineare inversa (6.5.24) ai poli "analogici" s_{pi}, cosicché il *modulo* dei poli z_{pi} si ottiene dalla (6.5.25). Se allora s_{p_i} ha parte reale negativa (cioè se $\sigma < 0$), il modulo del corrispondente polo z_{p_i} è *minore di uno*, e il sistema a tempo discreto è a sua volta stabile. In termini di trasformazione tra piani complessi, il *semipiano a parte reale negativa* $\sigma < 0$ del piano s si trasforma nel *cerchio di raggio unitario* con centro nell'origine $|z| < 1$ sul piano z (Figura 6.48).

La risposta in frequenza $\overline{H}(f)$ del sistema a tempo discreto ottenuto tramite trasformazione bilineare si ottiene come di consueto valutando la funzione di trasferimento sui punti della circonferenza di raggio unitario $|z| = 1$:

$$\overline{H}(f) = H\left(e^{j2\pi fT}\right) \tag{6.5.26}$$

Vediamo di calcolare quali sono i punti del piano s che nella corrispondenza (6.5.23) generano i punti della circonferenza $|z| = 1$. Sostituiamo quindi $z = e^{j2\pi fT}$ nell'espressione della trasformazione bilineare, ottenendo

$$s = \frac{2}{T} \cdot \frac{1 - e^{-j2\pi fT}}{1 + e^{-j2\pi fT}} = \frac{2}{T} \cdot \frac{e^{-j\pi fT}}{e^{-j\pi fT}} \frac{e^{j\pi fT} - e^{-j\pi fT}}{e^{j\pi fT} + e^{-j\pi fT}} = \frac{2}{T} \cdot \frac{j\sin(\pi fT)}{\cos(\pi fT)} = j\frac{2}{T}\tan(\pi fT) \tag{6.5.27}$$

ovvero, indicando con f_a la frequenza relativa al piano s del sistema a tempo continuo, si ha

$$\sigma + j2\pi f_a = j\frac{2}{T}\tan(\pi fT) \tag{6.5.28}$$

Questa relazione è verificata per i punti del piano s con $\sigma = 0$, cioè *giacenti sull'asse dei numeri immaginari*, per i quali

$$2\pi f_a = \frac{2}{T}\tan(\pi fT) \tag{6.5.29}$$

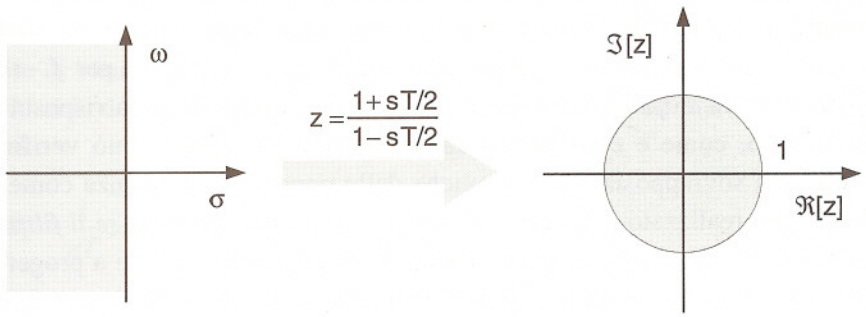


Figura 6.48 Stabilità e trasformazione bilineare

ovvero

$$f_a = \frac{1}{\pi T}\tan(\pi f T) \tag{6.5.30}$$

Questa relazione stabilisce una corrispondenza *diretta* tra i valori delle *risposte in frequenza* $H_a(f_a)$ del filtro analogico e $\overline{H}(f)$ del filtro a tempo discreto:

$$\overline{H}(f) = H_a\left(\frac{1}{\pi T}\tan(\pi f T)\right) \tag{6.5.31}$$

Ciò significa che le caratteristiche *frequenziali* (in particolare la *selettività*) del filtro realizzato mediante la trasformazione bilineare sono unicamente determinate dalle caratteristiche frequenziali del prototipo analogico. Confrontiamo questa relazione con quella relativa alla tecnica dell'invarianza impulsiva:

$$\overline{H}(f) = \sum_{k=-\infty}^{+\infty} H_a\left(f - \frac{k}{T}\right) \tag{6.5.32}$$

La prima osservazione è che la trasformazione bilineare *non dà luogo ad aliasing*. Infatti, l'andamento della risposta in frequenza del filtro a tempo discreto si può ricavare da quella del prototipo analogico attraverso il grafico di Figura 6.49 (che rappresenta il legame (6.5.30)) e la relazione (6.5.31). I valori della risposta in frequenza analogica $H(f)$ con f che varia da $-\infty$ a $+\infty$ sono riportati ("mappati") sull'intervallo frequenziale base $[-1/2T, 1/2T]$ della risposta del filtro numerico $\overline{H}(f)$, secondo la curva del tipo "arcotangente" di Figura 6.49. La scala delle frequenze subisce cioè una *compressione* nonlineare (chiamata *distorsione*, o con il termine anglosassone *warping*) nel passaggio dal sistema

analogico al sistema a tempo discreto. In particolare, il valore della risposta in frequenza analogica alla frequenza nulla viene conservato (la curva di distorsione passa per l'origine), e i valori-limite della risposta analogica per $f \to \pm\infty$ vengono riportati rispettivamente alle frequenze $f = \pm 1/2T$ sulla risposta del filtro discreto, come è evidenziato dalla Figura 6.50. Non si può verificare quindi alcuna sovrapposizione di repliche della risposta in frequenza come nel caso del filtro realizzato attraverso invarianza impulsiva, *neanche se il filtro ha banda illimitata*. La trasformazione bilineare è allora adatta anche a progettare filtri di tipo passa-alto o elimina-banda non realizzabili attraverso il metodo dell'invarianza impulsiva.

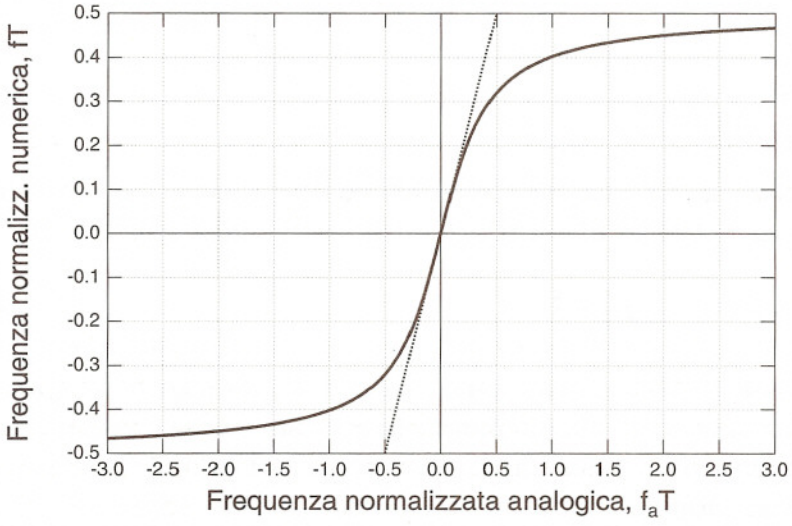


Figura 6.49 Relazione tra frequenza del sistema analogico f_a e frequenza f del corrispondente sistema a tempo discreto

Osserviamo inoltre che la relazione di distorsione dell'asse della frequenza di Figura 6.49 è *monotona crescente*. Per questa ragione la trasformazione bilineare *conserva le proprietà di selettività del prototipo analogico* utilizzato nel progetto: se, ad esempio, si vuole progettare un filtro passa-alto a tempo discreto, si deve far riferimento a un prototipo analogico passa-alto. L'unico inconveniente del quale occorre tener conto nell'applicazione di questa tecnica di progetto è la *deformazione* della risposta in frequenza conseguente alla nonlinearità della curva di distorsione.

Supponiamo adesso di voler progettare un filtro a tempo discreto con limite di

banda B assegnato. Affinché la specifica venga rispettata, il prototipo analogico che costituisce la base del progetto non deve essere caratterizzato dallo *stesso* limite di banda, bensì da un limite di banda B_a artificialmente *incrementato* in modo che, dopo la trasformazione bilineare, si ottenga effettivamente il limite voluto:

$$B_a = \frac{1}{\pi T}\tan(\pi B T) \tag{6.5.33}$$

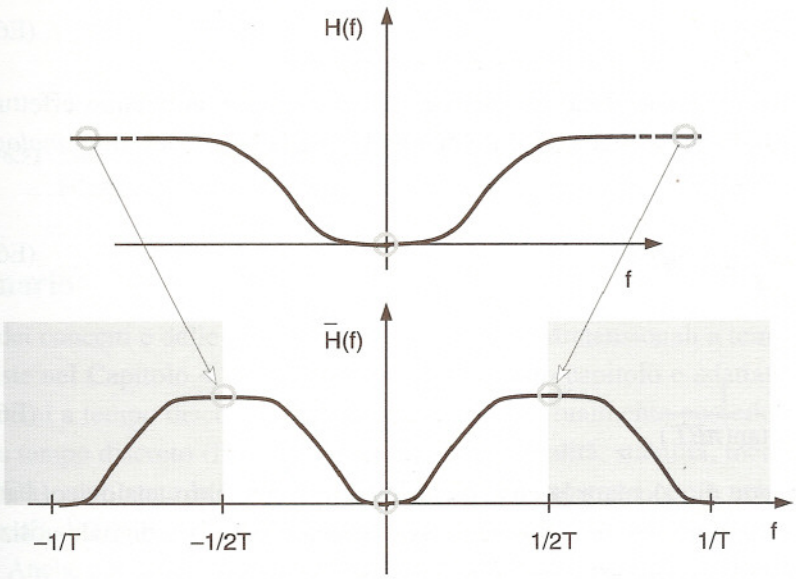


Figura 6.50 Compressione della risposta in frequenza nella trasformazione bilineare

L'operazione di incremento del limite di banda del prototipo viene chiamata *predistorsione* (o *prewarping*). La predistorsione permette di *precompensare* l'effetto della distorsione dell'asse delle frequenze nel passaggio dal sistema analogico a quello a tempo discreto, e consente quindi di rispettare la specifica di banda del filtro numericò.

Esempio 6.20

Si vuol progettare con la tecnica della trasformazione bilineare un filtro passa-basso digitale utilizzando come prototipo il filtro *R-C* degli Esempi 6.18-6.19. In questo caso però, a differenza di quanto visto con la tecnica dell'invarianza impulsiva, assegnamo una specifica di progetto richiedendo che il filtro a tempo

discreto (per il quale è assegnata la frequenza di campionamento $1/T$) abbia una banda a –3 dB pari a B. Come già visto negli esempi precedenti, la funzione di trasferimento del sistema analogico è data da

$$H_a(s)=\frac{1}{1+\alpha s} \tag{E6.20.1}$$

ove $\alpha = RC$ è la costante di tempo del circuito. La banda a –3dB B_a del filtro analogico è dunque

$$B_a=\frac{1}{2\pi\alpha} \tag{E6.20.2}$$

Per rispettare la specifica del *filtro a tempo discreto*, dobbiamo effettuare la *predistorsione* della banda, cioè dobbiamo fissare la banda del filtro *analogico* al valore

$$B_a=\frac{1}{\pi T}\tan(\pi BT) \tag{E6.20.3}$$

da cui

$$\frac{\alpha}{T}=\frac{1}{2\tan(\pi BT)} \tag{E6.20.4}$$

che permette di calcolare la costante di tempo α del filtro analogico. Fatto ciò possiamo determinare la funzione di trasferimento del filtro digitale utilizzando la trasformazione bilineare (6.5.22):

$$H(z)=H_a\left(\frac{2}{T}\frac{1-z^{-1}}{1+z^{-1}}\right)=\frac{1}{1+\alpha\dfrac{2}{T}\dfrac{1-z^{-1}}{1+z^{-1}}}=\frac{1}{1+\dfrac{1}{\tan(\pi BT)}\dfrac{1-z^{-1}}{1+z^{-1}}} \tag{E6.20.5}$$

da cui si ricava

$$H(z)=\left(\frac{\tan(\pi BT)}{\tan(\pi BT)+1}\right)\cdot\frac{1+z^{-1}}{1+\left(\dfrac{\tan(\pi BT)-1}{\tan(\pi BT)+1}\right)z^{-1}} \tag{E6.20.6}$$

La realizzazione in forma canonica del filtro a tempo discreto è rappresentata in Figura 6.51, mentre le risposte in ampiezza/fase sono riportate in Figura 6.52 per il caso particolare di $B=1/5T$. Confrontiamo tali risposte con quelle di Figura

6.46 relative allo *stesso* filtro realizzato con la tecnica dell'invarianza impulsiva: si nota l'assenza di aliasing, ma anche una distorsione della risposta in ampiezza/fase rispetto a quella del prototipo analogico. Naturalmente, la banda a –3 dB del filtro discreto è *esattamente* quella richiesta.

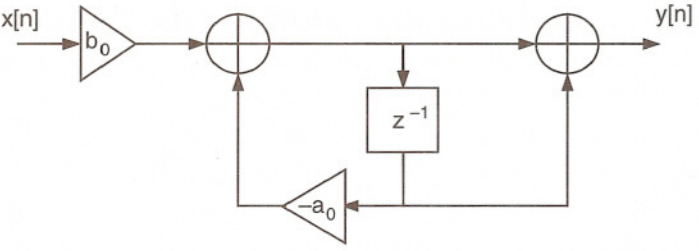


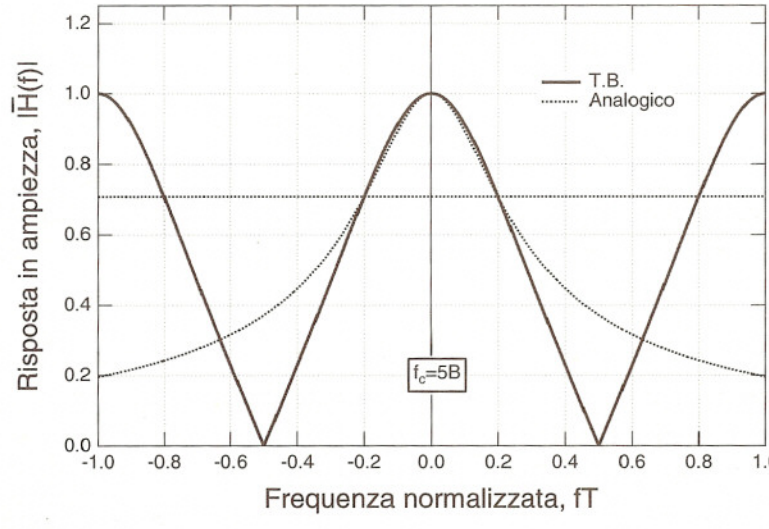
Figura 6.51 Struttura canonica del filtro a tempo discreto realizzato tramite trasformazione bilineare: $b_0 = \tan(\pi BT)/(\tan(\pi BT)+1)$, $a_0 = (\tan(\pi BT)-1)/(\tan(\pi BT)+1)$

Sommario

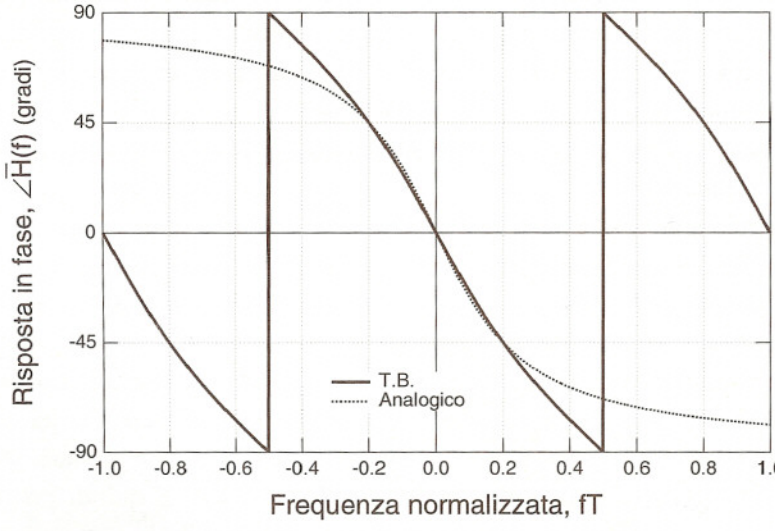
Molti dei concetti e delle proprietà dei sistemi monodimensionali a tempo continuo viste nel Capitolo 4 sono state riprese in questo capitolo e adattate al caso dei sistemi a tempo discreto. Tutte le proprietà eventualmente possedute dai sistemi a tempo discreto (linearità, stazionarietà, causalità, stabilità, memoria, invertibilità) sono state brevemente discusse, e si è mostrato come non vi siano sostanziali differenze rispetto al caso a tempo continuo, salvo piccoli adattamenti. Anche per il caso dei sistemi lineari e stazionari a tempo discreto valgono le stesse considerazioni: i concetti di risposta impulsiva, risposta in frequenza, convoluzione ecc. sono analoghi ai corrispondenti a tempo continuo, e sono stati brevemente richiamati.

Due sistemi caratteristici dell'elaborazione dei segnali a tempo discreto sono viceversa il *sovracampionatore* e il *decimatore*. Il sovracampionatore è costituito dalla cascata di due sistemi: il primo aumenta di un fattore M la frequenza di campionamento della sequenza d'ingresso inserendo $M-1$ campioni nulli tra due campioni consecutivi; il secondo è un filtro *interpolatore* digitale. Lo spettro della sequenza sovracampionata è *identico* a quello della sequenza di partenza, ma la sua *periodicità* è differente, poiché la frequenza di campionamento di uscita f_c' è maggiore di un fattore M: $f_c' = Mf_c = M/T$. È quindi possibile compiere operazioni di filtraggio *numerico* su questo spettro, in particolare eliminare *numericamente* le $M-1$ immagini di un segnale analogico campionato

nell'intervallo di frequenze $[-M/(2T), M/(2T)]$. Viceversa, il *decimatore* è costituito da un sistema che preleva un campione di segnale ogni M, riducendo così di un fattore M la frequenza di campionamento, e deve essere preceduto da un *filtro passa-basso* numerico anti-aliasing di banda $1/(2MT)$. Lo spettro della sequenza decimata è infatti una somma di M repliche dello spettro originario, ciascuna riportata a cavallo della frequenza $k/(MT)$, $k = 0,1,\ldots,M-1$.



(a)



(b)

Figura 6.52 Risposte in ampiezza (a) e fase (b) del filtro di Figura 6.51

Molto spesso è conveniente studiare i segnali e i sistemi lineari stazionari (SLS) a tempo discreto attraverso la *trasformata Z*. Quest'ultima rappresenta l'analogo a tempo discreto della trasformata di Laplace a tempo continuo, e permette di trattare anche segnali a energia illimitata. Attraverso la trasformata $\mathcal{Z}$ in particolare si analizzano e sintetizzano facilmente i SLS a tempo discreto descrivibili attraverso *equazioni alle differenze* lineari a coefficienti costanti. Di questi sistemi si può dare una rappresentazione grafica *diretta* o *canonica* che è basata sui soli tre operatori *amplificatore*, *ritardatore*, *sommatore*, e che aiuta molto nell'implementazione dei medesimi. La funzione di trasferimento del sistema $H(z)$, cioè la trasformata $\mathcal{Z}$ della risposta impulsiva, è una funzione razionale fratta in z^{-1}, e questo permette di verificare alcune proprietà del sistema (ad esempio la stabilità) molto facilmente. Sulla base della forma della $H(z)$ i SLS descritti da equazioni alle differenze si possono inoltre suddividere in due grandi classi: sistemi FIR a risposta impulsiva di durata finita e IIR a risposta impulsiva di durata infinita.

I sistemi IIR possono essere sintetizzati (progettati) con una certa facilità se si ricorre a uno tra i due criteri dell'*invarianza impulsiva* o della *trasformazione bilineare*. In entrambi i casi si parte da un filtro *prototipo* analogico di caratteristiche note. Il passaggio al filtro numerico avviene nel primo caso imponendo che la risposta impulsiva del filtro numerico si ottenga dal *campionamento* della risposta (continua) del filtro prototipo. Nel secondo caso, la funzione di trasferimento $H(z)$ del filtro numerico si ottiene trasformando la funzione di trasferimento $H(s)$ del filtro prototipo secondo la relazione $s = (2/T)(1-z^{-1})/(1+z^{-1})$. L'invarianza impulsiva provoca aliasing nella risposta in frequenza del filtro numerico, ed è quindi utilizzabile solo per filtri in cui l'aliasing possa essere annullato o reso trascurabile (non è perciò applicabile per filtri passa-alto o elimina-banda). Viceversa, la trasformazione bilineare può essere usata per la sintesi di qualsiasi tipo di filtro, ma può comportare (specialmente per piccole frequenze di campionamento) sensibili distorsioni della risposta del filtro numerico rispetto a quella del prototipo analogico.

Esercizi proposti

6.1 Un sistema a tempo discreto ha risposta in frequenza

$$\overline{H}(f) = \left[2\sin^2(\pi fT) + \cos(4\pi fT)\right] e^{-j4\pi fT}$$

Disegnare la realizzazione canonica di tale sistema e determinarne la risposta $g[n]$ al gradino unitario $u[n]$.

6.2 Utilizzando la tecnica dell'invarianza impulsiva, progettare il simulatore a tempo discreto di un filtro analogico con risposta in frequenza

$$H(f) = \cos^2\left(\frac{\pi f}{2B}\right) \text{rect}\left(\frac{f}{2B}\right)$$

nei due casi in cui la frequenza di campionamento è pari a:

a) $f_s = 1/T = 2B$;
b) $f_s = 1/T = B$

Determinare inoltre in entrambi i casi la risposta $y[n]$ del simulatore a tempo discreto alla sequenza $x[n] = u[n] - u[-n]$

6.3 Determinare la sequenza $y[n]$ in uscita al filtro a tempo discreto avente risposta impulsiva

$$h[n] = \frac{u[n] - u[n-8]}{8}$$

sapendo che la sequenza in ingresso a tale filtro è

$$x[n] = n\,,\ n = 0,1,...,7\ \ ,\ \ x[n+8] = x[n]\ \ \forall n$$

6.4 Si desidera progettare un filtro passa-basso a tempo discreto a partire dal prototipo di filtro a tempo continuo mostrato in Figura 6.53 (si ponga $RC = \alpha$). Il filtro a tempo discreto viene realizzato con una frequenza di campionamento $f_s = 1/T = 20$ kHz, e la sua banda a –40 dB deve essere $B = 5$ kHz. Rappresentare la struttura canonica di tale filtro e ricavarne la risposta impulsiva $h[n]$.

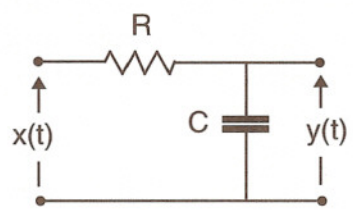


Figura 6.53

6.5 Si vuole progettare un simulatore a tempo discreto del sistema a tempo continuo di Figura 6.54. Scegliere a ragion veduta fra la tecnica della trasformazione bilineare e la quella dell'invarianza impulsiva. Con la tecnica

scelta, determinare poi la funzione di trasferimento $H(z)$ del simulatore discreto sapendo che la sua banda a –3 dB deve essere $B = 8$ kHz con una frequenza di campionamento $f_s = 1/T = 48$ kHz, e disegnarne infine la struttura canonica.

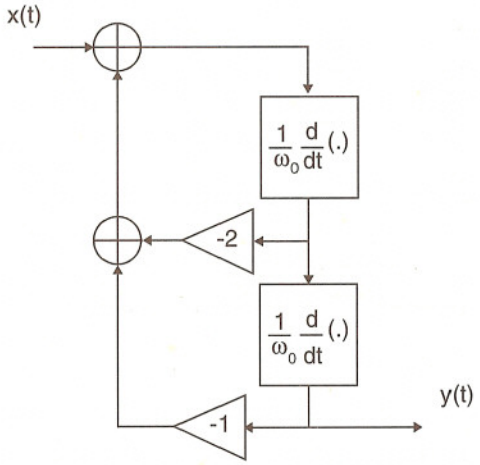


Figura 6.54

6.6 Nel sistema di Figura 6.55 la frequenza di campionamento è $f_s = 48$ kHz; il segnale di ingresso è inoltre $x(t) = 2 + \sin(\pi f_0 t)$ con $f_0 = 24$ kHz. Il sistema nonlineare (NL) a tempo discreto è caratterizzato dalla relazione ingresso-uscita $z[n] = x^2[n]$. Progettare un filtro FIR $h[n]$ causale a 3 coefficienti in modo tale che la sequenza di uscita sia $y[n] = \sin(n\pi/2)$.

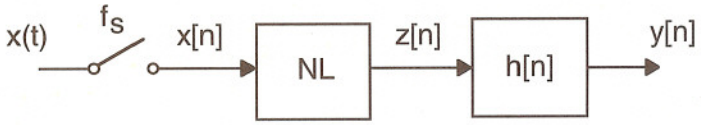


Figura 6.55

6.7 Progettare il simulatore a tempo discreto del sistema a tempo continuo $H(s)$ raffigurato in Figura 6.56, imponendo che la *frequenza di notch* f_N del sistema $H(z)$ così ottenuto sia pari a un quarto della frequenza di campionamento del segnale. Determinare inoltre il valore del parametro a in modo che il filtro a tempo discreto risultante (di cui si deve dare rappresentazione canonica) sia di tipo FIR.

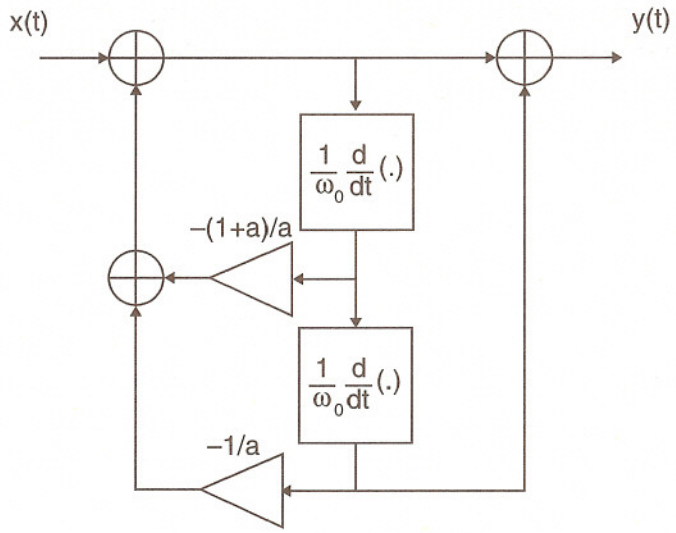


Figura 6.56

6.8 Considerando il sistema lineare invariante a tempo continuo di Figura 6.57, determinare la funzione di trasferimento e la risposta impulsiva $h[n]$ del relativo simulatore a tempo discreto ottenuto mediante il metodo della trasformazione bilineare (porre $\alpha = L/R$). Rappresentare inoltre la forma canonica del simulatore.

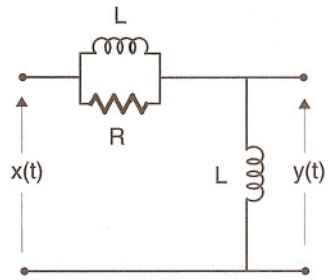


Figura 6.57

6.9 Il segnale

$$x(t) = \left(1 - \frac{|t|}{T_0}\right) \text{rect}\left(\frac{t}{2T_0}\right)$$

viene campionato con la frequenza di campionamento $f_c = 1/T = 4/T_0$. La

sequenza $x[n]$ risultante viene poi periodicizzata con periodo $N_0 = 4$, ottenendo la sequenza periodica $y[n]$ che viene filtrata con un filtro $h[n]$ ottenuto attraverso trasformazione bilineare del circuito rappresentato in Figura 6.58, con $RC = T$. Determinare l'espressione del segnale $z[n]$ in uscita da tale filtro.

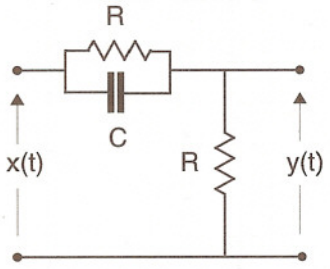


Figura 6.58

6.10 Calcolare la funzione di trasferimento $H(z)$ e la risposta in frequenza $\overline{H}(f)$ del SLS a tempo discreto avente la risposta impulsiva $h[n]$ illustrata in Figura 6.59, e dire poi se tale sistema è

a) causale o meno;

b) FIR o IIR;

c) passa-basso o passa-alto;

d) stabile o instabile.

Determinarne infine la risposta al gradino $g[n]$.

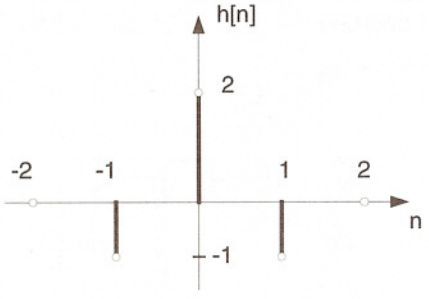


Figura 6.59

6.11 Progettare il simulatore a tempo discreto del sistema analogico rappresentato in Figura 6.60 attraverso il metodo dell'invarianza impulsiva, sapendo che la frequenza di campionamento è $f_c = 1/T = 4/T_0$. Determinare quindi

la risposta $y[n]$ del simulatore alla sequenza $x[n]$ rappresentata in Figura 6.61.

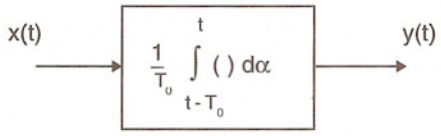


Figura 6.60

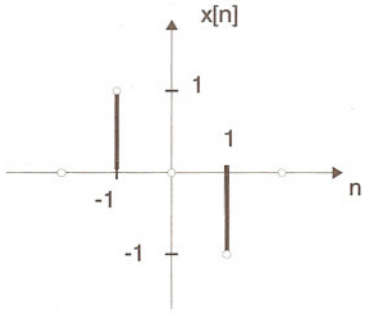


Figura 6.61

6.12 Progettare un filtro elimina-banda a tempo discreto con *frequenza di notch* $F_N = 12$ kHz e frequenza di campionamento $f_s = 1/T = 48$ kHz. A tal fine utilizzare il prototipo analogico di Figura 6.62, ponendo per comodità $f_0 = 1/\left(2\pi\sqrt{LC}\right)$ e sapendo che $\pi f_0 L/R = 1$. Rappresentare la forma canonica del filtro discreto e determinare la sequenza di uscita $y[n]$ quando in ingresso è posta l'eccitazione $x[n] = 2 + 2\sin^2(2\pi n f_1 T)$, $f_1 = 6$ kHz.

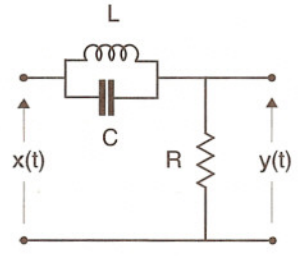


Figura 6.62

6.13 Determinare la risposta $y[n]$ del sistema a tempo discreto rappresentato in Figura 6.63 all'eccitazione

$$x[n] = 2 + \sin\left(\frac{\pi n}{2}\right) + \cos(\pi n)$$

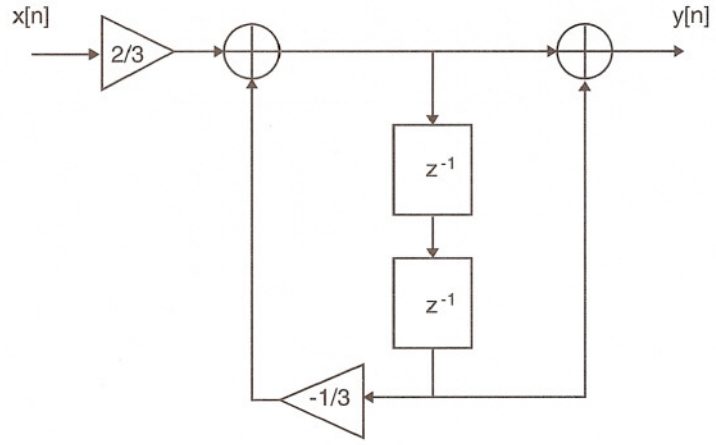


Figura 6.63

6.14 Si consideri il sistema a tempo discreto mostrato in Figura 6.64, dove $a_1 = 4/\left(4+\sqrt{6}\right)$, $a_2 = 4\sqrt{6}/\left(4+\sqrt{6}\right)$ e $K = 1/\left(4+\sqrt{6}\right)$. Si chiede di:

a) determinare l'equazione alle differenze del sistema;

b) stabilire se il filtro è passa-basso;

c) determinare la risposta $y[n]$ alla sequenza di ingresso $x[n] = 4 + (-1)^n$.

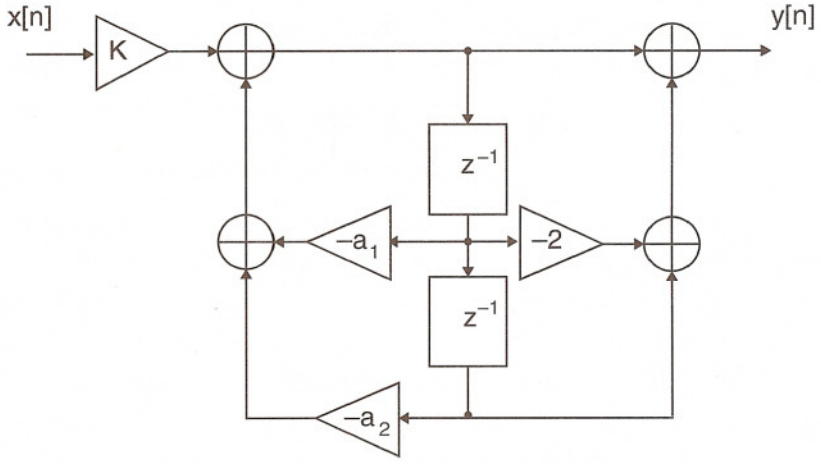


Figura 6.64

6.15 Considerando il sistema a tempo continuo di Figura 6.65, determinare la forma canonica e la risposta impulsiva $h[n]$ del relativo simulatore a tempo discreto ottenuto mediante il metodo della trasformazione bilineare nel caso $\omega_0 = 2/T$.

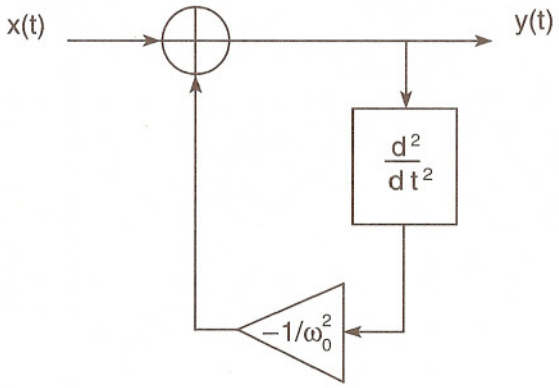


Figura 6.65

6.16 Indicare con $g(t)$ la risposta al gradino del sistema a tempo continuo di Figura 6.66. Determinare la funzione di trasferimento $H(z)$ del simulatore discreto la cui risposta al gradino $g[n]$ è data da $g[n] = g(nT)$ e rappresentarne la struttura canonica.

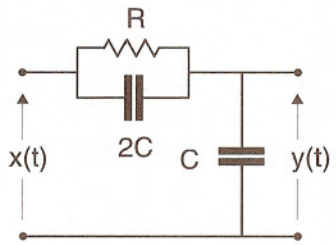


Figura 6.66

7

Richiami di teoria della probabilità

Premessa

Dobbiamo momentaneamente abbandonare lo studio dei segnali in senso stretto per richiamare alcuni concetti matematici fondamentali alla comprensione della natura e delle particolari proprietà dei segnali aleatori. Questi ultimi infatti vengono studiati attraverso gli strumenti della *teoria della probabilità e delle variabili aleatorie*, che si presuppone già nota al lettore. Lo scopo di questo capitolo è soltanto quello di *richiamare* i risultati principali della teoria della probabilità, e di riformularli con le notazioni che saranno poi riprese nel prossimo capitolo. Il lettore che si ritiene ferrato su questi argomenti può usare questo capitolo soltanto come *riferimento*, e può proseguire lo studio dei segnali procedendo direttamente alla lettura del Capitolo 8.

7.1 Esperimenti deterministici e aleatori

Ogni volta che si devono compiere misurazioni per controllare il verificarsi di certi avvenimenti, è necessario effettuare un *esperimento*. La definizione di esperimento che possiamo dare, così come potrebbe trovarsi su di un vocabolario, è appunto quella di una *prova pratica intesa alla verifica di una certa ipotesi di lavoro.* Supponiamo che si desideri studiare la "caduta di un grave", cioè di un corpo materiale, lasciato libero a una certa altezza h dal suolo, e in particolare che si desideri verificare la formula galileiana del tempo di caduta $t_0 = \sqrt{2h/g}$ (g =accelerazione di gravità). Dobbiamo costruire un *esperimento* che consiste nell'effettuare una *prova* di caduta, misurando con la massima accuratezza pos-

sibile il tempo impiegato per arrivare al suolo. I dati raccolti in molte prove permetteranno quindi di *verificare l'ipotesi* che sta alla base dell'*esperimento* stesso.

Dobbiamo però notare una importante diversità di fondo tra certi tipi di esperimenti e altri, che porta alla introduzione di due diverse *categorie* di esperimenti. Tornando sull'esempio precedente di caduta di un grave, e supponendo di effettuare molte prove, notiamo che i *risultati* che otteniamo in ogni prova sono molto simili fra loro. Nel limite in cui l'effetto della resistenza dell'aria risulta trascurabile o può essere a sua volta calcolato (ad esempio, conducendo l'esperimento in ambiente controllato al chiuso e usando corpi sferici costruiti con materiale ad alta densità) otteniamo risultati praticamente *identici* di volta in volta. Possiamo allora dire che l'esperimento condotto è di carattere *deterministico*, nel senso che è possibile *prevederne* il risultato *a priori*, cioè prima di effettuare una prova dell'esperimento stesso. Quest'ultimo mostra dunque una completa predicibilità: esiste una *legge* di carattere matematico che rende di fatto inutile la materiale effettuazione di una prova perché ne predice accuratamente il risultato.

Sorvoliamo ovviamente sulle implicazioni che questo concetto di determinismo ha sulla visione del mondo e della vita umana, e procediamo presentando un diverso tipo di esperimento. Immaginiamo dunque di sederci alla cassa di un supermercato e di contare il numero di clienti che si presentano nell'intervallo di tempo di un'ora dalle ore 11.00 alle ore 12.00 di ogni giorno. Questo può ritenersi a buon diritto un *esperimento* in virtù del quale è possibile verificare il numero di casse che sono necessarie per evitare lunghe code. Sfortunatamente, questo esperimento dà risultati anche molto diversi da prova a prova, e non è possibile *prevedere* a priori il risultato di nessuna delle prove effettuate. Il numero di persone osservato di volta in volta cambia in maniera *aleatoria*[1], cioè casuale. Bisogna dunque rinunciare a descrivere questo esperimento con una legge? La risposta fortunatamente è *no*. Ovviamente, sarà comunque impossibile trovare una legge che potrà *predire* in ogni prova il numero di clienti osservato: questo dato è ricavabile solo *a posteriori*, cioè *dopo* che la prova stessa è stata effettuata. Quel che è possibile fare in questo caso però è predire il comportamento *globale* dei dati che si ottengono *effettuando molte prove dell'esperimento*. In quest'ultimo caso, infatti, i dati raccolti mostrano quella che viene

[1] L'aggettivo *aleatorio* deriva dal latino *alea*, cioè dado, che è ritenuto l'oggetto casuale per antonomasia.

chiamata *regolarità statistica*. Nell'esperimento aleatorio per antonomasia, il lancio di un dado, nessuno è in grado di predire il risultato di una data prova, e cioè la faccia che si presenta lanciando il dado a un certo istante. L'esperienza però suggerisce che se abbiamo la pazienza di effettuare molti lanci del dado, diciamo 6000 (possibilmente con un apposito meccanismo!), osserveremo all'incirca 1000 volte la faccia 1, all'incirca 1000 volte la faccia 2,..., all'incirca 1000 volte la faccia 6. Questa regolarità permette di ricavare *anche per l'esperimento aleatorio* alcune leggi cui l'esperimento ottempera, però nel senso statistico appena accennato.

Come l'analisi matematica tradizionale è lo strumento matematico per eccellenza che descrive gli esperimenti deterministici (si pensi alla relazione tra l'analisi infinitesimale e la dinamica dei corpi), così la *teoria della probabilità* è lo strumento matematico sviluppato appositamente per descrivere gli esperimenti aleatori. Quello che un tempo veniva chiamato "calcolo delle probabilità" nasce infatti tra il diciassettesimo e il diciottesimo secolo a opera principalmente del matematico svizzero J. Bernoulli e dei matematici francesi B. Pascal e P.S. de Laplace per quantificare le vincite di giocatori e gestori dei giochi d'azzardo (dadi, carte, estrazioni di palline ecc.). E a giudicare dai guadagni dei casinò in tutto il mondo, è palese che questa teoria funziona alquanto bene!

7.2 Elementi di teoria della probabilità

7.2.1 Esperimento aleatorio, spazio di probabilità e proprietà elementari

Immaginiamo dunque di effettuare un esperimento aleatorio. Vediamo come la teoria della probabilità permetta di *modellare* strettamente questo esperimento, e come si possano poi ricavare delle *leggi* applicabili all'esperimento stesso. Per caratterizzare tale esperimento dobbiamo individuare innanzitutto l'insieme di tutti i suoi possibili risultati (ad esempio, le possibili facce del dado, o il numero di clienti che si possono presentare in un'ora alla cassa): tale insieme è detto *spazio campione* e si indica, convenzionalmente, con la lettera Ω. Se l'esperimento prevede un numero finito (dado) o infinito numerabile (clienti alla cassa) di risultati, questi verranno indicati con il simbolo ω_i, i=1,.... Quindi, con la notazione tipica della teoria degli insiemi, possiamo scrivere $\Omega = \{\omega_1, \omega_2, \ldots\}$.

Oltre che i singoli risultati dell'esperimento, è spesso importante considerare anche dei *gruppi* di risultati. Ad esempio, nell'esperimento della cassa del supermercato è importante considerare tutti i casi in cui il numero di clienti in un'ora è maggiore (ad esempio) di 20, perché essi potrebbero rappresentare un

numero eccessivo per quella cassa. L'interesse nell'esperimento potrebbe essere allora concentrato su questo *evento*: "numero di clienti in un'ora superiore a 20", cioè su tutti quei risultati dello spazio campione contenuti nel *sottoinsieme* {21,22,...}. I gruppi di risultati dello spazio campione sono chiamati *eventi*. Formalmente, gli *eventi* sono tutti i sottoinsiemi dello spazio campione che soddisfano le seguenti condizioni:

- se A è un evento, anche il suo complemento $\overline{A}$, rispetto all'insieme Ω, è un evento;
- se A e B sono eventi, anche la loro unione $A \cup B$ è un evento.

Usando queste proprietà si può anche dimostrare che:

- l'intersezione $A \cap B$ di due eventi arbitrari A e B è un evento;
- dato un evento A, gli insiemi $A \cup \overline{A}$ e $A \cap \overline{A}$ sono eventi: il primo, coincidente con Ω (ovvero con tutto lo spazio campione), è detto *evento certo*, mentre il secondo, indicato con il simbolo $\varnothing$ e non contenente alcun risultato dell'esperimento, è detto *evento impossibile*.

Le definizioni e proprietà sopra elencate dicono che gli eventi di uno spazio campione costituiscono una *classe* S, ovvero un insieme *chiuso* rispetto alle operazioni di unione e di intersezione.

A questo punto possiamo introdurre la *caratterizzazione completa di un esperimento aleatorio* che richiede sostanzialmente tre elementi: i) la descrizione del suo *spazio campione* Ω; ii) l'individuazione della sua *classe degli eventi* S, e infine iii) la descrizione della sua *legge di probabilità* $\Pr\{\cdot\}$ che associa ad ogni evento una misura della sua probabilità di presentazione. La terna $(\Omega, S, \Pr\{\cdot\})$ che rappresenta la descrizione dell'esperimento è chiamata *spazio di probabilità*. Qualche volta, con libertà di linguaggio, indicheremo con *esperimento aleatorio* ciò che in realtà è lo *spazio di probabilità,* identificando cioè l'esperimento con la propria descrizione matematica astratta.

Non è il caso di esaminare qui le varie definizioni e interpretazioni della probabilità di un evento, che sono oggetto di discussione nei testi specificamente dedicati alla teoria della probabilità e alla statistica (si veda comunque a questo proposito l'Esempio 7.1). Secondo lo scopo di questo capitolo, ci limiteremo qui a richiamare le proprietà base della probabilità la cui conoscenza sarà indispensabile allo studio elementare dei processi aleatori nel Capitolo 8.

La probabilità per una classe di eventi può essere definita secondo la *teoria assiomatica* la cui forma moderna si può sostanzialmente far risalire al matematico russo A.N. Kolmogorov. Secondo questo approccio, assegnato un esperi-

mento aleatorio con uno spazio campione Ω e la relativa classe degli eventi S, una legge di probabilità $\Pr(\cdot)$ è semplicemente una *corrispondenza* che associa a ogni elemento di S, cioè a ogni *evento* di interesse in una prova dell'esperimento, un numero reale che soddisfa i seguenti *assiomi*:

A.1 - la probabilità di un evento arbitrario A è non negativa:

$$\Pr(A) \geq 0 \tag{7.2.1}$$

A.2 - la probabilità dell'evento certo è unitaria (*assioma di normalizzazione*):

$$\Pr(\Omega) = 1 \tag{7.2.2}$$

A.3 - dati due eventi A e B mutuamente esclusivi (o incompatibili, o disgiunti, cioè che non possono verificarsi contemporaneamente), la probabilità dell'evento unione è data dalla somma delle probabilità di A e B:

$$A \cap B = \varnothing \Rightarrow \Pr(A \cup B) = \Pr(A) + \Pr(B) \tag{7.2.3}$$

Da questi *assiomi* si possono poi ricavare alcune *proprietà* (cioè dimostrare alcuni *teoremi* o corollari) che sembrano ovvie, ma che devono comunque essere ricondotte ai soli princìpi primi (cioè agli assiomi stessi):

- Dato un evento A, la probabilità dell'evento complementare $\overline{A}$ è data dal complemento a uno di $\Pr(A)$:

$$\Pr(\overline{A}) = 1 - \Pr(A) \tag{7.2.4}$$

- L'insieme impossibile ha probabilità nulla di verificarsi;
- La probabilità di un evento A non può assumere un valore maggiore di uno:

$$0 \leq \Pr(A) \leq 1 \tag{7.2.5}$$

- Dati due eventi A e B, la probabilità dell'evento unione $A \cup B$ è espressa dall'uguaglianza

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(B \cap A) \tag{7.2.6}$$

L'intersezione fra due eventi A e B può anche essere rappresentata con la scrittura AB, così come talvolta l'unione tra eventi viene indicata con la scrittura $A+B$. La probabilità dell'evento intersezione fra A e B $\Pr(B \cap A) = \Pr(AB)$, è chiamata *probabilità congiunta* degli eventi A e B. Le probabilità $\Pr(A)$ e $\Pr(B)$ sono dette, invece, *probabilità marginali*.

Data una coppia di eventi A e B, con $\Pr(B) \neq 0$, la probabilità $\Pr(A|B)$ dell'evento A *condizionata* al verificarsi dell'evento B è definita dalla relazione

$$\Pr(A|B) \triangleq \frac{\Pr(AB)}{\Pr(B)} \tag{7.2.7}$$

Nella teoria della probabilità $\Pr(A)$ è detta comunemente *probabilità a priori* dell'evento A e $\Pr(A|B)$ *probabilità a posteriori* di A dato B. La probabilità condizionata (in alcuni testi chiamata anche *condizionale*) ha un significato importante, che ruota attorno all'evento *condizionante* B. Infatti, $\Pr(A|B)$ è la probabilità che l'evento A assume *una volta che l'evento B si è già verificato*. La definizione (7.2.7) suggerisce proprio questo: la probabilità a priori di A viene scalata del fattore $1/\Pr(B)$ per tenere conto che l'evento B, essendosi già verificato, deve considerarsi come una sorta di "nuovo spazio campione" in quanto al di fuori di questo niente può verificarsi. In questo senso bisogna *rinormalizzare* tutte le probabilità rispetto a quella di B (si noti che $P(B \mid B) = 1$!).

Esempio 7.1 (Definizioni della probabilità di un evento)
Cerchiamo di definire un esperimento aleatorio (nel senso della definizione appena vista) che modelli il *lancio di un dado non truccato*. Evidentemente lo spazio campione Ω è costituito da 6 risultati $\{\omega_1, \omega_2, ..., \omega_6\}$, ove ω_i corrisponde al presentarsi al termine del lancio della faccia i-esima (cioè quella con un numero di puntini pari ad i). Poiché lo spazio campione Ω è *finito*, la classe degli eventi S è semplicemente costituita dalla collezione di tutti i sottoinsiemi di Ω stesso (che sono in numero, come è noto, di 2^6, inclusi $\varnothing$ ed Ω). La legge di probabilità degli eventi resta a questo punto definita non appena assegnamo una probabilità a ciascuno dei *risultati* dello spazio Ω. Fatto ciò, è possibile calcolare la probabilità di un qualunque evento A. Infatti, sfruttando la *simmetria* del problema, cioè l'ipotesi di dado non truccato, è ragionevole imporre che

$$\Pr(\omega_1) = \Pr(\omega_2) = \ldots = \Pr(\omega_6) = \frac{1}{6} \tag{E7.1.1}$$

Allora, ad esempio, la probabilità dell'evento A={*La faccia del dado è dispari*} è

$$\Pr(A) = \Pr(\{\omega_1\} \cup \{\omega_3\} \cup \{\omega_5\}) \tag{E7.1.2}$$

e quindi, per il terzo assioma,

$$\Pr(A) = \Pr(\{\omega_1\}) + \Pr(\{\omega_3\}) + \Pr(\{\omega_5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} \qquad \text{(E7.1.3)}$$

In casi come questo, di esperimento *simmetrico* e spazio campione finito, $\Pr(A)$ può anche essere calcolata attraverso la cosiddetta *definizione classica di probabilità,* attribuibile a Laplace. Questa prevede di individuare il numero $N_F(A)$ dei cosiddetti *casi favorevoli* ad A, e il numero N_P dei cosiddetti *casi possibili*. Quest'ultimo è semplicemente il numero totale di risultati contenuti in Ω, mentre il primo è il numero dei risultati elementari contenuti in A stesso. La probabilità cercata è allora data dal rapporto

$$\Pr(A) = \frac{N_F(A)}{N_P} \qquad \text{(E7.1.4)}$$

Considerando di nuovo $A = \{La\ faccia\ del\ dado\ è\ dispari\}$, è chiaro che $N_F(A) = 3$ ed $N_P = 6$, e quindi, dalla (E7.1.4), $\Pr(A) = 1/2$.

L'ipotesi cruciale alla base della definizione "classica" è la *perfetta simmetria* del dado o, in altri termini, l'*equiprobabilità* di tutti i possibili risultati dell'esperimento. Questa definizione ha il pregio di essere molto semplice, ma può essere applicata solamente a una classe ristretta di esperimenti. In particolare, essa è incapace di modellare il caso in cui il dado risulti "truccato", cioè le varie facce (intenzionalmente o per imperfezioni di manifattura) *non* siano equiprobabili!

In questi casi è più conveniente dare un'altra definizione di probabilità, che trova una sua giustificazione in un *approccio di tipo sperimentale*. Consideriamo di nuovo l'esperimento del lancio del dado ed effettuiamo N *prove* dell'esperimento stesso (cioè lanciamo il dado N volte). Indichiamo poi con N_A il numero di volte in cui, nelle suddette ripetizioni, si verifica l'evento A (cioè la faccia uscita mostra un numero dispari). All'aumentare di N, ovvero del numero di lanci, si ottiene una situazione simile a quella riassunta nella Tabella 7.1. Si nota una certa *regolarità* nella relazione tra il numero di volte in cui si verifica l'evento A rispetto al numero totale di prove effettuate: il rapporto N_A/N, detto *frequenza relativa dell'evento* A, approssima il numero 1/2 al crescere di N. Da questa osservazione discende la *definizione di probabilità di Von Mises* (o *frequentista*), secondo la quale

$$\Pr(A) = \lim_{N\to\infty} \frac{N_A}{N} \qquad \text{(E7.1.5)}$$

FREQUENZA RELATIVA DELL'EVENTO A

Tale definizione, rispetto alla definizione di Laplace, ha il vantaggio di

prescindere dalla simmetria del problema e di poter modellare anche il caso del dado "truccato", ma contiene un'operazione di limite che non si è in grado di eseguire (e che pone anche questioni di esistenza).

Tabella 7.1 Prove di lancio di un dado

N	N_A
100	47
1000	491
10000	4984
100000	50012
...	...

È interessante osservare che la definizione "frequentista" non è in contrasto con quella assiomatica di Kolmogorov. Infatti la probabilità $\Pr(A)$, espressa dalla frequenza relativa (E7.1.5), è i) una quantità *non negativa* poiché prodotta dal limite di un rapporto fra quantità positive; se inoltre ii) l'evento A coincide con l'evento certo Ω, allora banalmente $N_A = N$ e quindi $\Pr(A) = 1$; se infine iii) A e B sono due eventi mutuamente esclusivi, una prova dell'esperimento che fa verificare $A \cup B$ dà un risultato che sta o in A o in B, ma che non può stare in entrambi, per cui $N_{A\cup B} = N_A + N_B$ e allora

$$\Pr(A \cup B) = \lim_{N\to\infty} \frac{N_{A\cup B}}{N} = \lim_{N\to\infty} \frac{N_A + N_B}{N}$$
$$= \lim_{N\to\infty} \frac{N_A}{N} + \lim_{N\to\infty} \frac{N_A}{N} = \Pr(A) + \Pr(B) \qquad \text{(E7.1.7)}$$

cioè tutti gli assiomi di Kolmogorov sono automaticamente verificati! ❑

Due eventi A e B sono *indipendenti* se la probabilità marginale $\Pr(A)$ e la probabilità condizionata $\Pr(A|B)$ sono identiche, cioè in pratica se il verificarsi dell'evento B non ha alcuna influenza sull'evento A:

$$\Pr(A) = \Pr(A|B) \qquad (7.2.8)$$

o, tenendo conto della definizione di probabilità condizionata (7.2.7):

$$\Pr(AB) = \Pr(A) \cdot \Pr(B) \qquad (7.2.9)$$

Pertanto, se gli eventi A e B sono indipendenti, la probabilità congiunta $\Pr(AB)$ è pari al prodotto delle probabilità marginali. La relazione (7.2.9) viene utilizzata spesso per definire l'indipendenza fra due eventi, anche se il suo

significato non è immediatamente comprensibile come per la (7.2.8).

Consideriamo ora la coppia di eventi A e B, ciascuno avente probabilità non nulla. La probabilità condizionata $\Pr(A|B)$ può essere ricavata con la relazione

$$\Pr(A|B) = \frac{\Pr(B|A)\Pr(A)}{\Pr(B)} \tag{7.2.10}$$

nota come **teorema (o formula) di Bayes**. Se l'evento A è indipendente dall'evento B, cioè se la (7.2.8) è verificata, dalla (7.2.10) si ricava immediatamente l'uguaglianza $\Pr(B|A) = \Pr(B)$. La relazione di indipendenza tra due eventi gode, pertanto, della proprietà di simmetria.

Il teorema di Bayes è spesso usato insieme al *teorema della probabilità totale* che esaminiamo di seguito. Costruiamo una *partizione* dello spazio Ω scegliendo N eventi B_i, $i = 1, 2, \ldots, N$ di S con le seguenti proprietà (si veda la Figura 7.1):

$$B_i \cap B_k = \varnothing \quad \text{se} \quad i \neq k \tag{7.2.11}$$

$$\bigcup_{i=1}^{N} B_i = \Omega \tag{7.2.12}$$

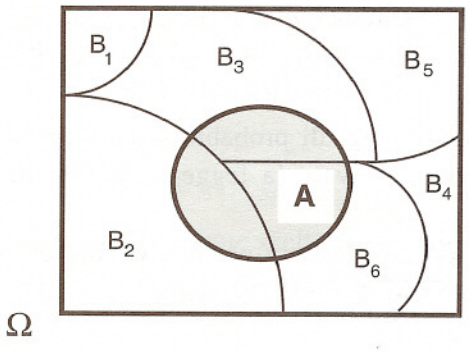


Figura 7.1 Dimostrazione grafica del teorema della probabilità totale

La probabilità di un evento A può allora essere calcolata (si ricordi il terzo assioma A.3) come segue:

$$\Pr(A) = \Pr(A \cap \Omega) = \Pr\left(A \cap \bigcup_{i=1}^{N} B_i\right) = \Pr\left(\bigcup_{i=1}^{N} A \cap B_i\right) = \sum_{i=1}^{N} \Pr(A \cap B_i) \tag{7.2.13}$$

Se in questa relazione si esprime poi ciascuna probabilità congiunta $\Pr(A \cap B_i)$ come prodotto tra la probabilità condizionata $\Pr(A|B_i)$ e la probabilità marginale $\Pr(B_i)$, si ricava la seguente uguaglianza:

$$\Pr(A) = \sum_{i=1}^{N} \Pr(A|B_i) \Pr(B_i) \tag{7.2.14}$$

che rappresenta appunto l'enunciato del **teorema della probabilità totale**.

7.2.2 Esperimento aleatorio composto

Consideriamo ora due *diversi* esperimenti aleatori caratterizzati dagli spazi campione Ω_1 e Ω_2 (ad esempio il lancio di un dado e l'estrazione di una carta da un mazzo di 52). È possibile definire un *esperimento composto* i cui risultati sono costituiti da una *coppia ordinata* dei risultati degli esperimenti componenti (ad esempio, (⚅,2♠)). Lo spazio campione Ω dell'esperimento composto è costituito dal *prodotto cartesiano* degli spazi dei due esperimenti componenti, cioè $\Omega = \Omega_1 \times \Omega_2$. Consideriamo, adesso, un evento A_1 definito nello spazio campione Ω_1 e un evento A_2 definito in Ω_2; vogliamo calcolare la probabilità dell'evento $A = A_1 \times A_2$ appartenente allo spazio campione Ω. Se i due esperimenti sono indipendenti, la probabilità dell'evento composto A è

$$\Pr(A) = \Pr_1(A_1) \cdot \Pr_2(A_2) \tag{7.2.15}$$

ove $\Pr_1(\cdot)$ e $\Pr_2(\cdot)$ rappresentano le leggi di probabilità definite rispettivamente per il primo e il secondo esperimento componente. È importante osservare che

- dalla conoscenza delle leggi di probabilità dei singoli esperimenti non è possibile, in generale, ricavare la legge di probabilità dell'esperimento composto;
- le considerazioni appena illustrate per una coppia di esperimenti aleatori indipendenti possono essere immediatamente estese al caso di n esperimenti aleatori indipendenti.

In questo ambito è utile ricordare il problema delle *prove ripetute binarie e indipendenti* (o *prove di Bernoulli*). In tal caso l'esperimento composto è costituito da n esperimenti identici, indipendenti, e aventi ciascuno uno spazio campione costituito da due risultati soltanto. In questo modello ricadono numerosi esperimenti elementari (testa/croce, vero/falso, 0/1, alto/basso ecc.). Indichiamo con ω_0 e ω_1 questi risultati e con $p = \Pr\{\omega_0\}$ e $q = \Pr\{\omega_1\} = 1 - p$ le loro rispettive probabilità, e definiamo l'evento $A = \{\omega_0$ *si presenta k volte in n prove*

ripetute}. La *formula di Bernoulli* (o *binomiale)* dice che

$$\Pr(A)=\binom{n}{k}p^k q^{n-k} \quad , \quad 0\le k\le n \tag{7.2.16}$$

ove compare il coefficiente binomiale

$$\binom{n}{k}=\frac{n!}{k!(n-k)!} \tag{7.2.17}$$

Esempio 7.2

Una scatola contiene due monete: la prima è una moneta "perfetta", la seconda è una moneta "truccata" avente $\Pr\{Testa\}=0.8$. Viene scelta casualmente una delle due monete, che viene poi lanciata per dieci volte in condizioni indipendenti, osservando l'uscita di 5 facce Testa e 5 facce Croce. Sulla base di quest'ultima osservazione, qual è la probabilità che la moneta scelta sia quella "perfetta"? Cosa può dirsi di questa probabilità se si osservano 5000 facce Testa su 10000 lanci?

In questo esempio si deve calcolare la probabilità *condizionata* $\Pr(A|B)$, ove B è l'evento osservato {*5 facce Testa, 5 facce Croce*}, mentre A è l'evento {*La moneta scelta è quella perfetta*}. Usando il teorema di Bayes (7.2.10) è possibile ricavare $\Pr(A|B)$ dalla conoscenza delle probabilità $\Pr(B|A)$, $\Pr(A)$ e $\Pr(B)$. Poiché la scelta fra le due monete è casuale, abbiamo facilmente

$$\Pr(A)=\Pr(\overline{A})=\frac{1}{2} \tag{E7.2.1}$$

Supponiamo ora di aver effettivamente scelto la moneta perfetta, cioè che si sia verificato A. La probabilità condizionata $\Pr(B|A)$ può essere calcolata mediante la formula di Bernoulli immaginando che il risultato ω_0 corrisponda al presentarsi della faccia Testa, e ω_1 al presentarsi della faccia Croce. Applicando la (7.2.16) con $n=10$ lanci, $k=5$ uscite di Testa, $p=q=1/2$ (le facce della moneta perfetta sono equiprobabili), risulta

$$\Pr(B|A)=\binom{10}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^5=\binom{10}{5}\frac{1}{2^{10}}=\frac{252}{1024}\cong 0.246 \tag{E7.2.2}$$

$\Pr(B)$ può essere calcolata mediante il teorema della probabilità totale, con la partizione $\Omega=A\cup\overline{A}$:

$$\Pr(B) = \Pr(B|A)\Pr(A) + \Pr\left(B|\overline{A}\right)\Pr\left(\overline{A}\right) \tag{E7.2.3}$$

Resta da ricavare la probabilità $\Pr\left(B|\overline{A}\right)$ (cioè la probabilità di avere 5 volte Testa su 10 prove con la moneta *truccata*), che può di nuovo essere calcolata mediante la formula di Bernoulli con $n=10$, $k=5$, $p=\Pr\{\omega_0\}=0.8$ $q=\Pr\{\omega_1\}=0.2$ (le facce della moneta truccata non sono equiprobabili):

$$\Pr\left(B|\overline{A}\right) = \binom{10}{5}(0.8)^5(0.2)^5 = 252\cdot(0.16)^5 \cong 0.0264 \tag{E7.2.4}$$

Usando i valori appena ricavati possiamo finalmente determinare $\Pr(B)$:

$$\begin{aligned}\Pr(B) &= \binom{10}{5}\left(\frac{1}{2}\right)^{10}\cdot\frac{1}{2} + \binom{10}{5}(0.8)^5(0.2)^5\cdot\frac{1}{2} \\ &= \frac{252}{2}\left[\frac{1}{1024} + (0.16)^5\right] \cong 0.1363\end{aligned} \tag{E7.2.5}$$

e infine

$$\begin{aligned}\Pr(A|B) &= \frac{\Pr(B|A)\Pr(A)}{\Pr(B)} = \frac{\binom{10}{5}\left(\frac{1}{2}\right)^{2\cdot5}\cdot\frac{1}{2}}{\frac{1}{2}\binom{10}{5}\left(\frac{1}{2}\right)^{2\cdot5} + \frac{1}{2}\binom{10}{5}(0.8)^5(0.2)^5} \\ &= \frac{1}{1+(2\cdot0.8\cdot2\cdot0.2)^5} \cong 0.903\end{aligned} \tag{E7.2.6}$$

Questa probabilità è piuttosto alta. Lanciando infatti la moneta *truccata,* è abbastanza improbabile ottenere una serie di risultati "equiripartiti" 5-5 come quella richiesta.

Ripetendo i calcoli per il caso in cui B sia l'evento $\{$*5000 facce Testa, 5000 facce Croce*$\}$, si ottiene, in analogia alla (E7.2.6), la seguente espressione per $\Pr(A|B)$:

$$\Pr(A|B) = \frac{1}{1+(0.8\cdot0.2\cdot4)^{5000}} \cong 1 \tag{E7.2.7}$$

e si raggiunge quasi la *certezza* che la moneta scelta sia effettivamente quella perfetta. Infatti, la probabilità di ottenere un così preciso bilanciamento dei risultati $\{Testa\}$ e $\{Croce\}$ con la moneta (pesantemente) truccata è praticamente *nulla*. ❑

7.3 Variabili aleatorie

7.3.1 Definizione di variabile aleatoria

Consideriamo due diversi esperimenti aleatori. Il primo è il familiare lancio del dado (supposto perfetto) il cui spazio campione Ω_1 è noto. Il secondo consiste nello scegliere un qualunque giorno non festivo della settimana, ad esempio per verificare casualmente le presenze del personale di una ditta settimana per settimana. Lo spazio campione $\Omega_2 = \{Lun, Mar, Mer, Gio, Ven, Sab\}$ di questo secondo esperimento è evidentemente diverso da Ω_1. In realtà, i due esperimenti sono solo *apparentemente* diversi: in ogni esperimento si tratta di scegliere con completa casualità uno tra 6 possibili risultati distinti. Cerchiamo dunque di trovare una maniera di accomunare i due esperimenti, tentando di *astrarci* (cioè di allontanarci, di muoverci su di un livello maggiormente concettuale e più generale) dai due casi particolari. Potremmo ad esempio *numerare* le facce del dado e i giorni della settimana da 1 a 6 e considerare non più i particolari risultati dell'esperimento (diversi nei due casi), bensì il *valore* associato a ciascuno dei possibili *risultati*. Si ottiene in questo caso una "descrizione" che si attaglia a entrambi gli esperimenti, anche se in realtà essi si riferiscono a casi diversi con spazi campione diversi. In sintesi, abbiamo costruito una quantità numerica *variabile* (in questo caso da 1 a 6) a seconda del risultato dell'esperimento rispettivo, cioè un esempio di ciò che in teoria della probabilità viene chiamata *variabile aleatoria*.

Cerchiamo adesso di dare una definizione formale di variabile aleatoria più precisa di quella colloquiale appena vista. Consideriamo dunque un esperimento aleatorio avente uno spazio campione Ω, una classe degli eventi S e una legge di probabilità $\Pr(\cdot)$ (per semplicità considereremo uno spazio campione numerabile). Definiamo quindi (come sopra accennato) una *corrispondenza*, indicata con $X(\omega_i)$, che associa a ciascun risultato ω_i dell'esperimento un unico numero *reale*. Tale corrispondenza fra lo spazio Ω e l'asse reale è una *variabile aleatoria* se l'insieme di risultati dell'esperimento per i quali è verificata la disuguaglianza $X(\omega) \le a$ (tale insieme si indica convenzionalmente con la scrittura $\{X(\omega) \le a\}$) è un *evento*, comunque si scelga il valore del parametro reale a. Nelle pagine seguenti le variabili aleatorie saranno sempre rappresentate da lettere maiuscole omettendo la dipendenza dal risultato ω dell'esperimento (ad esempio X, Y, Z anziché $X(\omega)$, $Y(\omega)$, $Z(\omega)$ ecc.).

Il concetto di variabile aleatoria è utile ogniqualvolta si deve *modellare* un esperimento che ha come risultato finale un *valore numerico* non prevedibile a

priori. L'esempio per antonomasia è un'operazione di *misura* con uno strumento utilizzato ai limiti della precisione ottenibile. In questo caso, anche ripetendo più volte l'esperimento, si ottengono valori misurati di volta in volta diversi in maniera *aleatoria* per effetto dell'incertezza della misura. Questo procedimento è dunque immediatamente rappresentabile con una variabile aleatoria, con una grandezza cioè che varia in maniera non predicibile ogni volta che si ripete l'esperienza.

Dobbiamo adesso chiederci: una volta *definita* la variabile, come è possibile *trasferire* la legge di probabilità relativa alla classe S degli eventi dello spazio campione a insiemi di *valori* assunti dalla variabile aleatoria stessa? In un problema di misura, su cui è definita una variabile aleatoria X, è spesso significativo calcolare la probabilità che i valori di tale variabile aleatoria si trovino compresi in un intervallo del tipo $a < X \leq b$, che è di interesse per la misura stessa. Concettualmente, data la corrispondenza istituita (si veda la Figura 7.2), questo equivarrebbe a "ripercorrere la strada all'indietro" e in particolare a identificare in Ω tutti e soli quei risultati che danno valori della variabile aleatoria compresi appunto tra a e b. Questo insieme di risultati, per come è stata costruita la variabile aleatoria, è sicuramente un *evento* e a esso è associabile una *probabilità*. Quindi, con estensione di linguaggio, chiameremo direttamente *evento* anche l'intervallo di valori sulla retta reale $(a,b]$ perché a esso è comunque associabile una probabilità che indicheremo con $\Pr\{a < X \leq b\}$.

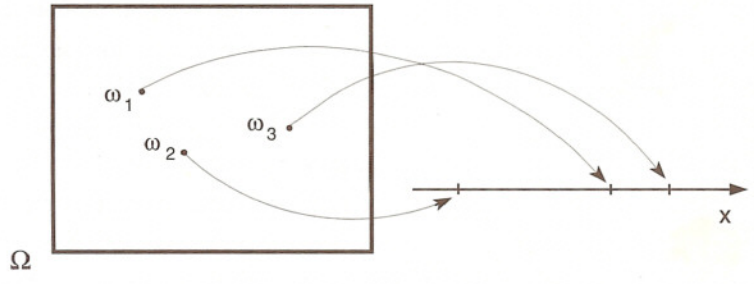


Figura 7.2 Definizione di una variabile aleatoria

L'operazione concettuale di "tracciamento a ritroso" appena descritta diventa superflua se si introduce la *funzione distribuzione di probabilità* $F_X(x)$ (talvolta chiamata anche funzione di *ripartizione*) di una variabile aleatoria X, definita come segue:

$$F_X(x) \triangleq \Pr\{X \le x\} \tag{7.3.1}$$

In questa definizione, x è un valore reale generico ma *fissato* (una sorta di valore di "sonda") che identifica l'evento $\{X \le x\}$ di cui si deve calcolare la probabilità. La funzione distribuzione gode delle seguenti proprietà fondamentali:

D.1 – essa assume valori appartenenti all'intervallo $[0, 1]$, ovvero

$$0 \le F_X(x) \le 1 \tag{7.3.2}$$

D.2 – il suo valore limite per $x \to +\infty$ vale 1, cioè

$$\lim_{x\to\infty} F_X(x) = F_X(+\infty) = P\{X \le +\infty\} = 1 \tag{7.3.3}$$

D.3 – il suo valore limite per $x \to -\infty$ vale 0, cioè

$$\lim_{x\to-\infty} F_X(x) = F_X(-\infty) = P\{X \le -\infty\} = 0 \tag{7.3.4}$$

D.4 – essa è monotona non decrescente, ovvero

$$x_2 > x_1 \;\Rightarrow\; F_X(x_2) \ge F_X(x_1) \tag{7.3.5}$$

D.5 – essa è continua da destra, cioè

$$F_X(x) = \lim_{h\to 0^+} F_X(x+h) \tag{7.3.6}$$

D.6 – se essa presenta una discontinuità di prima specie nel punto $x = \bar{x}$, allora la differenza fra il suo limite destro e il suo limite sinistro in tale punto è pari alla probabilità dell'evento $\{X = \bar{x}\}$, ovvero

$$\Pr\{X = \bar{x}\} = F_X(\bar{x}^+) - F_X(\bar{x}^-) \tag{7.3.7}$$

D.7 – la probabilità dell'evento $\{a < X \le b\}$ può essere calcolata mediante la relazione

$$\Pr\{a < X \le b\} = F_X(b) - F_X(a) \tag{7.3.8}$$

A seconda delle caratteristiche della funzione distribuzione di probabilità, le variabili aleatorie possono essere suddivise in tre classi: variabili *discrete*, *continue* e *miste*. Una variabile aleatoria X è *discreta* se la sua distribuzione $F_X(x)$ è una funzione costante a tratti, cioè è del tipo

$$F_X(x) = \sum_k \Pr\{X = x_k\}\, u(x - x_k) \tag{7.3.9}$$

Tenendo conto delle proprietà D.6 e D.7, ciò significa che la variabile aleatoria assume con probabilità diversa da zero un insieme di valori x_k *discreto* (finito o non). Le quantità $p_k = \Pr\{X = x_k\}$ sono le cosiddette *masse* di probabilità[2].

Se, invece, la distribuzione $F_X(x)$ è una funzione ovunque continua, allora la variabile aleatoria è *continua* e la probabilità che X assuma un valore reale fissato è identicamente nulla.

Una variabile aleatoria è *mista* se non appartiene a nessuna delle due classi ora definite. La funzione distribuzione di una variabile aleatoria mista è continua *quasi ovunque*, cioè continua salvo in un numero finito (o in una infinità numerabile) di punti. Esempi dell'andamento qualitativo di funzioni distribuzione di probabilità per una variabile aleatoria continua, discreta e mista sono mostrati in Figura 7.3.

7.3.2 Densità di probabilità di una variabile aleatoria

Il comportamento statistico di una variabile aleatoria X è dunque caratterizzato completamente dalla conoscenza della sua funzione distribuzione $F_X(x)$. Una descrizione alternativa è anche data dalla funzione *densità di probabilità* $f_X(x)$ della variabile, definita come segue:

$$f_X(x) \triangleq \frac{dF_X(x)}{dx} \tag{7.3.10}$$

Da questa definizione si ricava immediatamente la relazione inversa

$$F_X(x) = \int_{-\infty}^{x} f_X(u)du \tag{7.3.11}$$

La funzione densità di probabilità $f_X(x)$ gode delle seguenti proprietà:

- essa è non negativa, ovvero

  $$f_X(x) \geq 0 \tag{7.3.12a}$$

  per ogni valore reale della variabile x, poiché la funzione distribuzione è monotona non decrescente;

---

[2] Per rispettare le proprietà D.6 e D.7, la funzione gradino unitario che appare nella (7.3.9) deve essere definita in modo leggermente diverso da quanto visto nel Capitolo 3. In particolare, qui si deve intendere u(0)=1 (cioè si deve assicurare la continuità da destra).

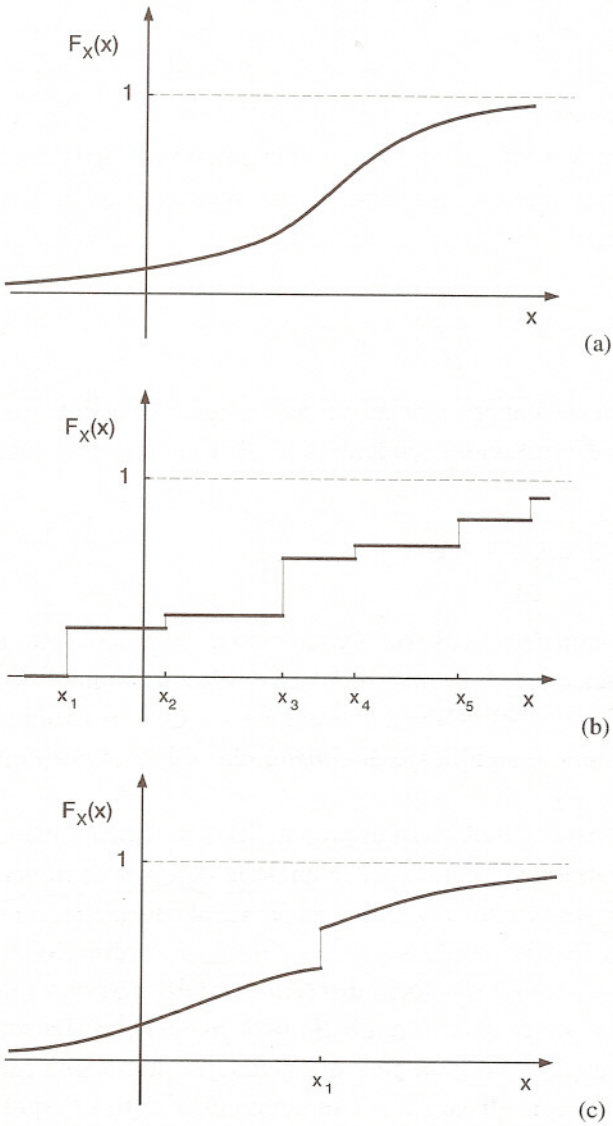


Figura 7.3 Esempio di funzioni distribuzioni di probabilità di una variabile aleatoria continua (a), discreta (b), mista (c)

- la probabilità dell'evento $\{a < X \leq b\}$ può essere calcolata con la relazione

$$\Pr\{a < X \leq b\} = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx \qquad (7.3.12b)$$

- l'integrale su tutto l'asse reale della funzione densità è unitario, cioè

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1 \tag{7.3.12c}$$

in quanto questo integrale rappresenta la probabilità dell'evento certo.

Consideriamo ora un intervallo di ampiezza molto piccola Δx attorno a un punto fissato $\bar{x}$. Si ha:

$$\Pr\{\bar{x} < X \le \bar{x} + \Delta x\} = \int_{\bar{x}}^{\bar{x}+\Delta x} f_X(x)\,dx \cong f_X(\bar{x}) \cdot \Delta x \tag{7.3.13}$$

ove evidentemente l'approssimazione dell'integrale è valida proprio nell'ipotesi di ampiezza dell'intervallo tendente a 0. Riorganizzando questa relazione si ottiene

$$f_X(\bar{x}) \cong \frac{\Pr\{\bar{x} < X \le \bar{x} + \Delta x\}}{\Delta x} \tag{7.3.14}$$

cioè il valore della densità di probabilità in un certo punto è pari al rapporto tra il contributo elementare di probabilità dato dall'intervallino considerato, e l'ampiezza (la *misura* cioè) dell'intervallino stesso. Questo risultato giustifica pienamente il nome e il significato di *densità* che abbiamo attribuito alla funzione $f_X(x)$

A questo riguardo, molti testi di probabilità e statistica trattano separatamente il caso della variabile *continua* per la quale la definizione di densità secondo la (7.3.10) è ben posta, e il caso di una variabile aleatoria *discreta* per la quale la definizione di densità perde significato in senso ordinario. La distribuzione $F_X(x)$ di una variabile aleatoria discreta è infatti *discontinua* e quindi non derivabile nei punti di discontinuità. Se però ripensiamo alla soluzione data nel Capitolo 3 al problema della derivata generalizzata di una funzione gradino, osserviamo che con gli strumenti matematici a nostra disposizione non c'è alcuna necessità di questa distinzione. Se infatti riprendiamo l'espressione (7.3.9) della funzione distribuzione di una variabile aleatoria discreta, ne ricaviamo immediatamente la funzione densità di probabilità:

$$f_X(x) = \sum_k p_k\,\delta(x - x_k) \tag{7.3.15}$$

che contiene naturalmente la funzione generalizzata δ di Dirac. Questo indica

ancora una volta che per una variabile aleatoria discreta la probabilità è *concentrata* nei particolari valori x_k dell'asse reale anziché essere *distribuita* con continuità come per una variabile aleatoria continua. In Figura 7.4 sono qualitativamente rappresentati alcuni esempi di densità di probabilità per una variabile continua, discreta e mista.

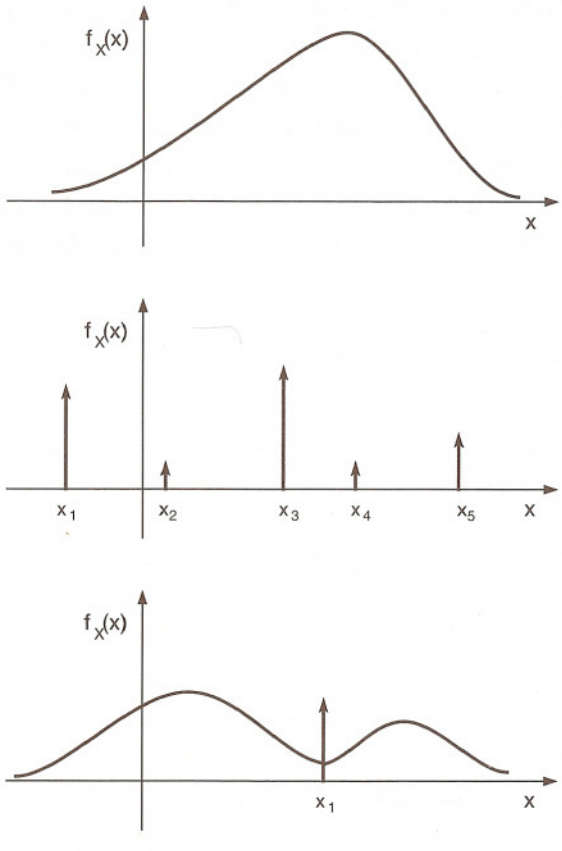


Figura 7.4 Esempio di funzioni densità di probabilità di una variabile aleatoria continua (a), discreta (b), mista (c)

Esempio 7.3

a) La variabile aleatoria uniforme

Una variabile aleatoria continua Y è *uniforme* sull'intervallo (a,b) se la sua densità di probabilità $f_Y(y)$ è costante in tale intervallo e si annulla al di fuori di esso. Per motivi di *normalizzazione* il valore della funzione su (a,b) deve essere evidentemente pari a $1/(b-a)$:

$$f_Y(y) = \frac{1}{b-a}\,\mathrm{rect}\left(\frac{y-(b+a)/2}{b-a}\right) \tag{E7.3.1}$$

Questo significa in pratica che Y non potrà mai assumere valori all'esterno di (a,b), mentre tutti i valori all'interno di tale intervallo vengono considerati alla stessa stregua. Infatti, ogni evento del tipo $\{\bar{x} < X \le \bar{x} + \Delta x\}$ contenuto in (a,b) ha sempre la medesima probabilità di verificarsi indipendentemente dalla posizione di $\bar{x}$. In un certo senso, la variabile aleatoria uniforme è l'analogo continuo della variabile aleatoria discreta definita su un insieme finito di valori equiprobabili (come nel caso del dado perfetto). L'insieme delle variabili aleatorie uniformi nell'intervallo (a,b) si denota comunemente con la scrittura $\mathcal{U}(a,b)$. Per indicare l'appartenenza della variabile aleatoria Y a questa classe si scrive $Y \in \mathcal{U}(a,b)$. La funzione distribuzione $F_Y(y)$ si ricava facilmente:

$$F_Y(y) = \int_{-\infty}^{y} f_Y(\alpha)\,d\alpha = \begin{cases} 0 & y < a \\ (y-a)/(b-a) & a \le y \le b \\ 1 & y > b \end{cases} \tag{E7.3.2}$$

L'andamento delle funzioni $f_Y(y)$ ed $F_Y(y)$ è illustrato nella Figura 7.5.

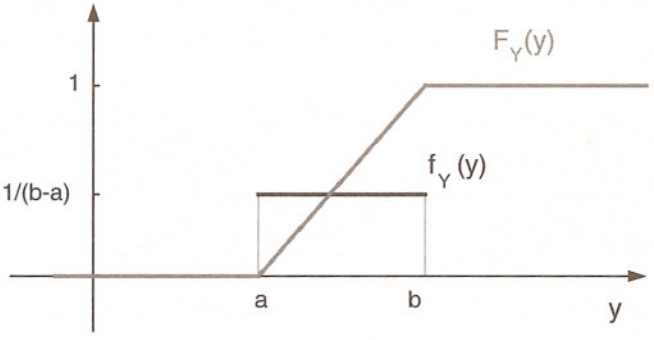


Figura 7.5 Funzioni densità e distribuzione di probabilità della variabile aleatoria uniforme

b) La variabile esponenziale

Una variabile aleatoria continua X è *esponenziale unilatera* se la sua densità di probabilità $f_X(x)$ è espressa dalla relazione

$$f_X(x) = \frac{1}{\eta}\exp\left(-\frac{x}{\eta}\right)\mathrm{u}(x) \tag{E7.3.3}$$

ove η è un parametro reale positivo il cui significato sarà chiaro in seguito. La variabile aleatoria esponenziale è usata comunemente in problemi di affidabilità e calcolo del rischio. In prima approssimazione, infatti, il cosiddetto *tempo di vita* di un dispositivo o di un apparato anche complesso, ovvero il tempo che intercorre prima che l'apparato vada per la prima volta fuori servizio, è modellabile con una variabile aleatoria esponenziale di parametro η opportuno. La sua funzione distribuzione $F_X(x)$ è allora

$$F_X(x) = \int_{-\infty}^{x} f_X(\alpha)d\alpha = \left[1 - \exp(-x/\eta)\right]\mathrm{u}(x) \qquad \text{(E7.3.4)}$$

L'andamento delle funzioni $f_X(x)$ ed $F_X(x)$ è illustrato nella Figura 7.6.

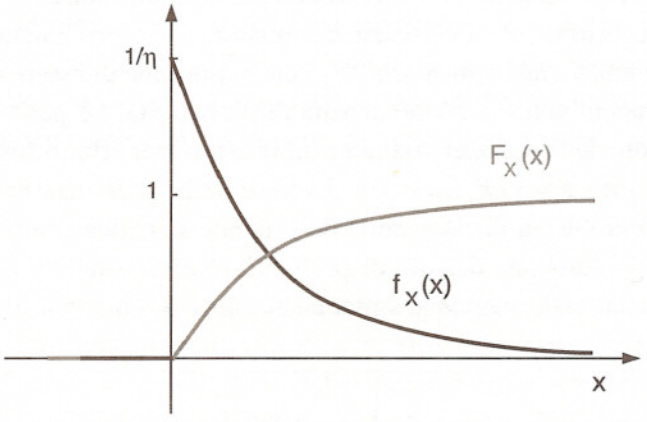


Figura 7.6 Funzioni densità e distribuzione di probabilità della variabile aleatoria esponenziale

c) La variabile di Poisson

Una variabile aleatoria discreta Z con densità di probabilità

$$f_Z(z) = \sum_{k=0}^{+\infty} e^{-\Lambda}\frac{\Lambda^k}{k!}\,\delta(z-k) \qquad \text{(E7.3.5)}$$

è *di Poisson* con parametro Λ ($\Lambda > 0$). Come si nota, la variabile Z può assumere con probabilità diversa da zero solo valori *interi non negativi*. L'insieme delle variabili aleatorie di Poisson con parametro Λ si indica, comunemente, con la scrittura $\mathcal{P}(\Lambda)$ e per denotare l'appartenenza della variabile aleatoria Z a questa classe si scrive $Z \in \mathcal{P}(\Lambda)$. La variabile aleatoria di Poisson è un buon modello dell'esperimento aleatorio di conteggio dei clienti descritto nel

Paragrafo 7.1, purché Λ venga scelto opportunamente. La funzione distribuzione $F_Z(z)$ è

$$F_Z(z) = e^{-\Lambda} \sum_{k=0}^{+\infty} \frac{\Lambda^k}{k!} \mathrm{u}(z-k) \qquad \text{(E7.3.6)}$$

L'andamento della massa di probabilità $p_k = e^{-\Lambda}\Lambda^k / k!$ della variabile di Poisson è infine mostrato nella Figura 7.7 per due diversi valori di Λ. ❑

7.3.3 Trasformazione di una variabile aleatoria

Nei problemi che coinvolgono una variabile aleatoria, è molto comune dover eseguire operazioni matematiche sui valori assunti dalla variabile stessa. Supponiamo di voler misurare il valore di una piccolissima corrente elettrica in un resistore. A causa dell'incertezza di misura, il valore misurato viene modellato come una variabile aleatoria X, di cui si presume di essere in grado di ricavare l'andamento della funzione densità di probabilità. Se però si desidera conoscere il valore della *potenza dissipata* sul resistore per effetto Joule, si deve calcolare la quantità $P = r \cdot X^2$, ove r è il valore della resistenza del resistore. Nasce dunque il problema di ricavare la descrizione statistica completa, e cioè l'andamento della funzione densità di probabilità, di questa *nuova* variabile aleatoria P ottenuta *trasformando* la variabile aleatoria originaria X.

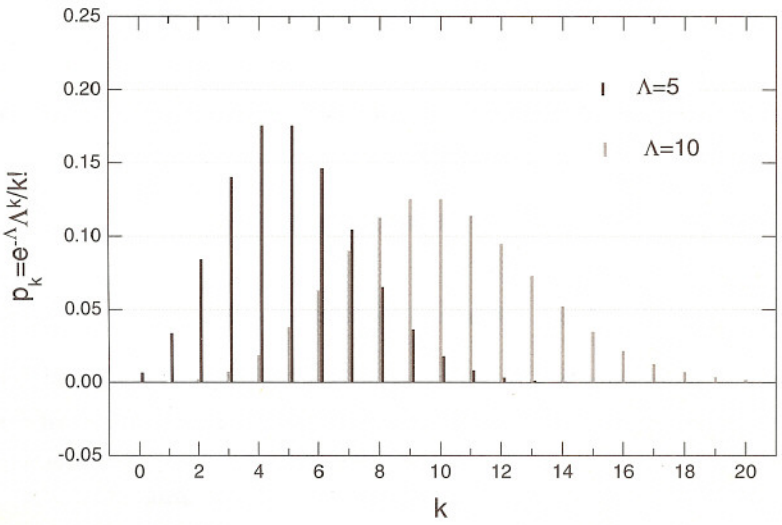


Figura 7.7 Massa di probabilità della variabile aleatoria di Poisson

Generalizziamo il problema: consideriamo una variabile aleatoria *continua* X a partire dalla quale viene definita una variabile aleatoria Y mediante la relazione

$$Y = g(X) \tag{7.3.16}$$

ove $g(x)$ è una funzione di variabile reale a valori reali. Nota la densità di probabilità $f_X(x)$ della variabile X, è possibile calcolare la densità di probabilità $f_Y(y)$ della variabile aleatoria Y mediante la relazione

$$f_Y(y) = \sum_i \frac{f_X(x_i)}{|g'(x_i)|} \tag{7.3.17}$$

ove l'insieme $\{x_i\}$ è costituito da tutte le soluzioni dell'equazione $g(x) = y$. Questo risultato è noto come **teorema fondamentale** per la trasformazione di una variabile aleatoria. Naturalmente, la dipendenza da y del secondo membro della (7.3.17) è "nascosta" nell'espressione degli x_i, il cui numero e valore dipende infatti dal particolare valore di y considerato. Per applicare correttamente il teorema fondamentale è allora importante avere ben presente quanto segue:

- a seconda del valore di y considerato, $\{x_i\}$ può essere un insieme vuoto (nel qual caso evidentemente $f_Y(y) = 0$) o può contenere un numero finito o infinito numerabile di punti;
- se nel punto $x = \bar{x}$ con $\bar{y} = g(\bar{x})$ la derivata prima $g'(\bar{x})$ è nulla si hanno due casi: i) la trasformazione $g(x)$ ha in $\bar{x}$ un massimo o un minimo relativi; se $f_X(\bar{x})$ è diverso da 0, allora la $f_Y(y)$ tenderà in $\bar{y}$ a $+\infty$; oppure ii) $\bar{x}$ appartiene a un intervallo I nel quale la funzione $g(x)$ assume un valore costante. In quest'ultimo caso, la variabile aleatoria Y assume il valore $\bar{y} = g(\bar{x})$ con probabilità

$$\Pr\{Y = \bar{y}\} = \Pr\{X \in I\} \tag{7.3.18}$$

e se tale probabilità è non nulla la variabile aleatoria Y è *mista.*

Esempio 7.4

Nel circuito elettrico di Figura 7.8 il generatore di tensione v_0 viene collegato alla squadra R-C all'istante $t = 0$. Il resistore r ha un *tempo di guasto* aleatorio X in corrispondenza del quale esso interrompe il circuito. L'istante X è una variabile aleatoria avente densità di probabilità esponenziale:

$$f_X(x) = \frac{1}{2\alpha} \exp\left(-\frac{x}{2\alpha}\right) u(x) \tag{E7.4.1}$$

con $\alpha = rc$. Vogliamo determinare la densità di probabilità $f_V(v)$ della variabile aleatoria V che rappresenta la tensione ai capi del condensatore dopo il guasto del resistore.

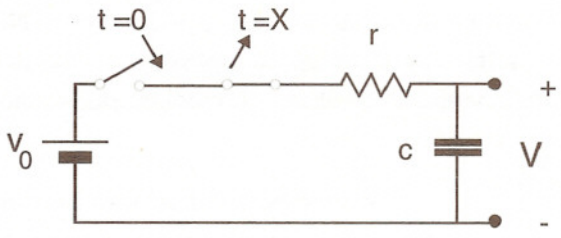


Figura 7.8 Schema della squadra *R-C* con guasto

L'andamento della tensione $v(t)$ ai capi del condensatore si trova dalla relazione di carica del condensatore c a partire dall'istante $t = 0$:

$$v(t) = v_0\left[1 - \exp(-t/\alpha)\right]\mathrm{u}(t) \tag{E7.4.2}$$

dalla quale segue immediatamente

$$V \equiv v(X) = v_0\left[1 - \exp(-X/\alpha)\right]\mathrm{u}(X) \tag{E7.4.3}$$

Abbiamo identificato una legge di *trasformazione* fra la variabile aleatoria X e la variabile aleatoria V, il cui grafico è rappresentato nella Figura 7.9. Cominciamo con l'osservare che l'equazione $v = g(x)$ non ha soluzioni se $v \geq v_0$ o $v < 0$; quindi la densità di probabilità $f_V(v)$ è *nulla* su questi intervalli. Resta quindi da determinare l'espressione di $f_V(v)$ per $0 \leq v < v_0$. Secondo il teorema fondamentale, bisogna per prima cosa trovare il numero e il valore dei punti x_i che soddisfano $v = v(x_i)$. Dalla Figura 7.9 è chiaro che nell'intervallo $[0, v_0)$ esiste sempre uno e un solo valore di x che soddisfa la relazione

$$v = v(x) \tag{E7.4.4}$$

e cioè

$$x_1 = -\alpha \ln\left(1 - \frac{v}{v_0}\right) \tag{E7.4.5}$$

Si ricava perciò:

$$f_V(v) = \frac{f_X(x_1)}{|v'(x_1)|} \tag{E7.4.6}$$

con

$$v'(x) = \frac{v_0}{\alpha}\exp(-x/\alpha) \tag{E7.4.7}$$

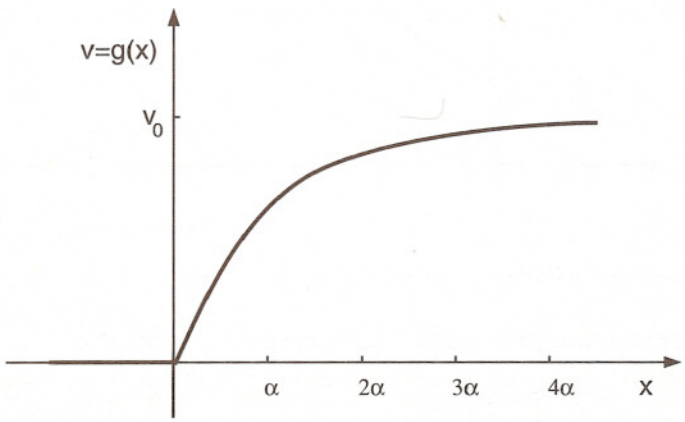


Figura 7.9 Legge di trasformazione della variabile aleatoria dell'Esempio 7.4

Riprendendo in considerazione la forma della densità di probabilità (E7.4.1) si ottiene infine

$$f_V(v) = \left.\frac{\frac{1}{2\alpha}\exp\left(-\frac{x}{2\alpha}\right)u(x)}{\left|\frac{v_0}{\alpha}\exp(-x/\alpha)\right|}\right|_{x=-\alpha\ln\left(1-\frac{v}{v_0}\right)} = \frac{\frac{1}{2\alpha}\sqrt{1-\frac{v}{v_0}}}{\frac{v_0}{\alpha}\left(1-\frac{v}{v_0}\right)} = \frac{1}{2v_0}\frac{1}{\sqrt{1-\frac{v}{v_0}}} \tag{E7.4.8}$$

per ogni $v \in [0, v_0)$. L'andamento della $f_V(v)$ è illustrato in Figura 7.10. ❑

7.3.4 Indici caratteristici di una distribuzione

La conoscenza della funzione densità (o distribuzione) di probabilità di una variabile aleatoria rappresenta il massimo di informazione che si può avere sul comportamento statistico dei valori assunti dalla variabile stessa. Naturalmente, però, non sempre è possibile arrivare a una conoscenza così completa riguardo a un problema aleatorio che si sta trattando. Molto più spesso, ci si accontenta della conoscenza di alcuni *parametri statistici semplificati* o *indici* relativi alla

distribuzione di probabilità presentata dalla variabile.

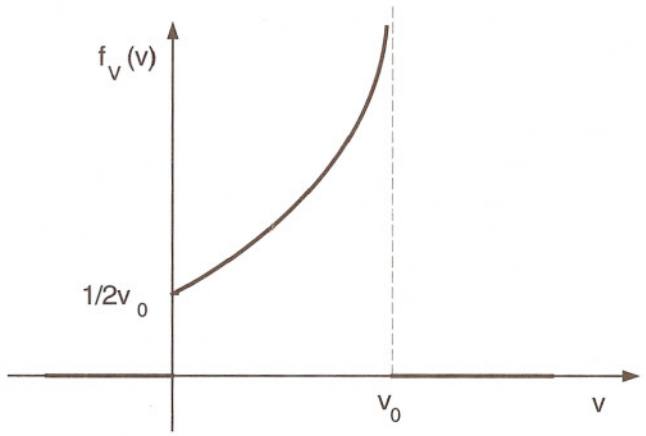


Figura 7.10 Densità di probabilità ricavata nell'Esempio 7.4

Il *valore atteso* (chiamato anche *valor medio, speranza, attesa*) η_X di una variabile aleatoria X con densità di probabilità $f_X(x)$ è definito dalla relazione

$$\eta_X \triangleq \int_{-\infty}^{\infty} x\, f_X(x)\, dx \tag{7.3.19}$$

e rappresenta in certo senso un valore "baricentrico" attorno al quale si distribuiscono i valori della variabile aleatoria stessa (indice di *posizione*). Se la variabile è discreta, richiamando la relazione (7.3.15) della relativa densità di probabilità, si ha

$$\eta_X \triangleq \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_{-\infty}^{\infty} x \sum_k p_k\, \delta(x - x_k)\, dx = \sum_k p_k \int_{-\infty}^{\infty} x\, \delta(x - x_k)\, dx$$
$$= \sum_k p_k\, x_k \tag{7.3.20}$$

che coincide con la formula normalmente fornita separatamente per le variabili discrete, e dove è stata sfruttata la *proprietà campionatrice* della funzione δ.

Quando si ha a che fare con un problema di *trasformazione* di una variabile aleatoria $Y = g(X)$, si introduce il cosiddetto *operatore valor medio*:

$$\mathrm{E}\{g(X)\} \triangleq \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx \tag{7.3.21}$$

La lettera E nell'operatore valor medio $E\{\cdot\}$ è l'iniziale della parola inglese *Expectation* che traduce l'italiano "aspettativa". Notiamo che anche il valore atteso η_X può essere riscritto formalmente usando questo operatore:

$$\eta_X = \int_{-\infty}^{\infty} x\, f_X(x)\, dx = E\{X\} \tag{7.3.22}$$

cioè il valore atteso è anche il valor medio della variabile aleatoria stessa! Se una variabile aleatoria Y è funzione della variabile aleatoria X, $Y = g(X)$, allora il valore atteso η_Y di Y, dato per definizione da

$$\eta_Y = \int_{-\infty}^{+\infty} y\, f_Y(y)\, dy \tag{7.3.23}$$

può essere calcolato con la relazione

$$\eta_Y = \mathrm{E}\{g(X)\} = \int_{-\infty}^{+\infty} g(x)\, f_X(x)\, dx \tag{7.3.24}$$

Questo risultato è noto come **teorema del valor medio**, ed è molto utile in pratica perché permette di evitare il calcolo esplicito della funzione $f_Y(y)$ che sarebbe necessario attraverso la definizione (7.3.23). Notiamo in margine che l'operatore valor medio, come è chiaro dalla sua stessa definizione attraverso un'operazione di integrazione, gode della proprietà di *linearità*:

$$\mathrm{E}\{\alpha \cdot g(X) + \beta \cdot h(X)\} = \alpha \cdot \mathrm{E}\{g(X)\} + \beta \cdot \mathrm{E}\{h(X)\} \tag{7.3.25}$$

con α e β costanti qualunque.

Due variabili aleatorie che pure presentano lo stesso valore atteso possono avere comportamenti statistici sensibilmente differenti. Nella Figura 7.11 sono mostrate appunto le densità di probabilità di due variabili aleatorie X_1 e X_2 accomunate dallo stesso valore atteso η_X, ma molto diverse per quel che riguarda la "larghezza" della relativa curva attorno al valore η_X. La densità $f_{X_1}(x)$ è molto "allargata" attorno al valore atteso; viceversa la $f_{X_2}(x)$ è molto più "appuntita". Questo suggerisce che per la prima variabile è piuttosto probabile trovare valori "lontani" dal valore atteso, e quindi molto più *dispersi* rispetto a quest'ultimo che per la variabile X_2. Per quantificare questo comportamento con un singolo parametro statistico si definisce la *varianza* σ_X^2 di una variabile aleatoria X come segue:

$$\sigma_X^2 \triangleq \mathrm{E}\left\{(X-\eta_X)^2\right\} = \int_{-\infty}^{\infty}(x-\eta_X)^2 f_X(x)\,dx \tag{7.3.26}$$

Il parametro σ_X, radice quadrata della varianza, è la cosiddetta *deviazione standard*. A maggiore varianza della variabile aleatoria corrispondono valori molto dispersi attorno al valor medio, e viceversa. Al limite, una variabile aleatoria che presenta varianza nulla ha valori *per niente* dispersi attorno al valor medio e la sua densità di probabilità diventa "infinitamente appuntita" attorno a questo valore:

$$f_X(x) = \delta(x-\eta_X) \tag{7.3.27}$$

In tal caso la probabilità di trovare valori diversi dal valore atteso è nulla, e quindi la variabile aleatoria in pratica "collassa" in un valore *certo*, precisamente nella quantità costante e non più variabile η_X.

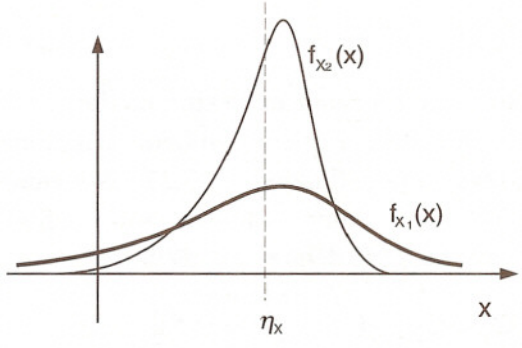


Figura 7.11 Densità di probabilità con ugual valore atteso ma diversa varianza

Il *valore quadratico medio* (talvolta chiamato anche *potenza*) di una variabile aleatoria è infine definito come

$$m_X^2 \triangleq \mathrm{E}\left\{X^2\right\} = \int_{-\infty}^{\infty} x^2\, f_X(x)\,dx \tag{7.3.28}$$

Poiché l'operatore valor medio $\mathrm{E}\{\cdot\}$ è un operatore *lineare*, è facile trovare il legame tra la varianza σ_X^2 e il valor quadratico medio m_X^2:

$$\sigma_X^2 \triangleq \mathrm{E}\left\{(X-\eta_X)^2\right\} = \mathrm{E}\left\{X^2+\eta_X^2-2\eta_X\cdot X\right\} = \mathrm{E}\left\{X^2\right\}+\mathrm{E}\left\{\eta_X^2\right\}-\mathrm{E}\left\{2\eta_X\cdot X\right\}$$
$$= m_X^2+\eta_X^2-2\eta_X^2 = m_X^2-\eta_X^2 \tag{7.3.29}$$

Esempio 7.5

Calcoliamo i parametri statistici semplificati per le variabili aleatorie considerate nell'Esempio 7.3. Il valore atteso η_X di una variabile aleatoria X *esponenziale unilatera*, avente densità di probabilità $f_X(x)$ espressa dalla (E7.3.3), è dato da (si veda la (7.3.19))

$$\eta_X = \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_0^{\infty} (x/\eta) \exp(-x/\eta)\, dx = \eta \tag{E7.5.1}$$

Dunque il parametro η, che abbiamo usato per caratterizzare la variabile aleatoria esponenziale unilatera, coincide in realtà con il *valore atteso* η_X della variabile aleatoria stessa. Il valor quadratico medio m_X^2 della variabile aleatoria esponenziale X è dato da

$$m_X^2 = \int_0^{\infty} (x^2/\eta) \exp(-x/\eta)\, dx = 2\eta_X^2 \tag{E7.5.2}$$

e, quindi, la sua varianza σ_X^2 è pari a

$$\sigma_X^2 = \mathrm{E}\{X^2\} - \eta_X^2 = \eta_X^2 = \eta^2 \tag{E7.5.3}$$

Lasciamo al lettore come esercizio la dimostrazione che il valore atteso, il valor quadratico medio e la varianza di una variabile aleatoria uniforme $Y \in \mathcal{U}(a,b)$ sono rispettivamente

$$\eta_Y = \frac{a+b}{2} \tag{E7.5.4}$$

$$m_Y^2 = \frac{a^2+ab+b^2}{2} \tag{E7.5.5}$$

$$\sigma_Y^2 = \frac{(b-a)^2}{12} \tag{E7.5.6}$$

Notiamo che all'aumentare dell'ampiezza del campo di variabilità $[a,b]$ per una variabile uniforme aumenta anche la *varianza* della variabile, come ci si deve attendere dal carattere di *indice di dispersione* che questo parametro riveste.

Consideriamo, infine, una variabile aleatoria discreta di Poisson $Z \in \mathcal{P}(\Lambda)$. Il suo valore atteso η_Z è dato da

$$\eta_Z = \sum_{k=0}^{+\infty} p_k z_k = e^{-\Lambda} \sum_{k=0}^{+\infty} e^{-\Lambda} \frac{\Lambda^k}{k!} \cdot k = e^{-\Lambda} \sum_{k=1}^{+\infty} \frac{\Lambda^k}{(k-1)!}$$

$$= \Lambda e^{-\Lambda} \sum_{k=0}^{+\infty} \frac{\Lambda^k}{k!} = \Lambda e^{-\Lambda} e^{\Lambda} = \Lambda \tag{E7.5.7}$$

Per determinare il valor quadratico medio e la varianza di Z, calcoliamo dapprima il seguente valor medio:

$$\mathrm{E}\{Z(Z-1)\} = \sum_{k=0}^{+\infty} p_k z_k (z_k - 1) = \sum_{k=0}^{+\infty} e^{-\Lambda} \frac{\Lambda^k}{k!} k(k-1) = e^{-\Lambda} \sum_{k=2}^{+\infty} \frac{\Lambda^k}{(k-2)!}$$

$$= \Lambda^2 e^{-\Lambda} \sum_{k=0}^{+\infty} \frac{\Lambda^k}{k!} = \Lambda^2 e^{-\Lambda} e^{\Lambda} = \Lambda^2 \tag{E7.5.8}$$

D'altronde,

$$\mathrm{E}\{Z(Z-1)\} = \mathrm{E}\{Z^2\} - E\{Z\} = m_X^2 - \eta_X = m_X^2 - \Lambda \tag{E7.5.9}$$

e quindi

$$m_X^2 = \Lambda^2 + \Lambda \tag{E7.5.10}$$

La varianza σ_Z^2 di Z è allora data da

$$\sigma_Z^2 = \mathrm{E}\{Z^2\} - \eta_Z^2 = \Lambda^2 + \Lambda - \Lambda^2 = \Lambda \tag{E7.5.11}$$

Dunque il parametro Λ, che caratterizza una variabile aleatoria di Poisson, ne rappresenta sia il valore atteso sia la varianza. ❑

7.3.5 La variabile aleatoria Gaussiana

Una variabile aleatoria X è *Gaussiana* o *normale* se la sua funzione densità di probabilità è

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-\eta_X)^2}{2\sigma_X^2}} \tag{7.3.30}$$

ove come di consueto σ_X^2 e η_X indicano rispettivamente la varianza e il valore atteso della variabile aleatoria stessa. Sinteticamente, scriveremo che $X \in \mathcal{N}(\eta_X, \sigma_X^2)$. In particolare, una variabile $N \in \mathcal{N}(0,1)$, ovvero una variabile aleatoria Gaussiana con valor medio nullo e varianza unitaria, è detta *variabile normale standard* e la sua densità di probabilità è

$$f_N(n) = \frac{1}{\sqrt{2\pi}} e^{-\frac{n^2}{2}} \tag{7.3.31}$$

il cui andamento è illustrato nella Figura 7.12. Usando il teorema fondamentale (7.3.17), si vede che una generica variabile $X \in \mathcal{N}(\eta_X, \sigma_X^2)$ può essere espressa come *trasformazione lineare* della variabile aleatoria normale standard N:

$$X = \sigma_X N + \eta_X \quad \Rightarrow \quad f_X(x) = \frac{1}{\sigma_X} f_N\left(\frac{x - \eta_X}{\sigma_X}\right) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{(x-\eta_X)^2}{2\sigma_X^2}} \tag{7.3.32}$$

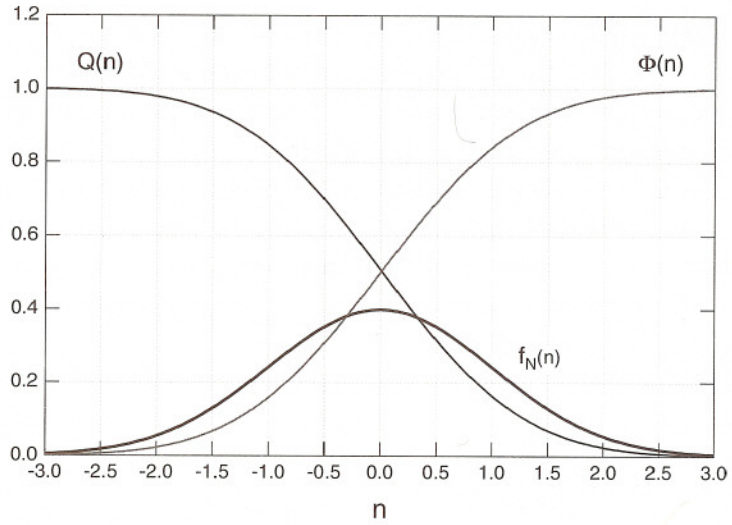


Figura 7.12 Grafico della densità di probabilità normale standard $f_N(n)$ e delle funzioni $\Phi(n)$ e $Q(n)$

La funzione distribuzione di una variabile aleatoria Gaussiana non può essere espressa in forma chiusa. Per questa ragione si definisce la funzione distribuzione $\Phi(x)$ per una variabile aleatoria normale standard:

$$\Phi(x) \triangleq F_N(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-n^2/2}\, dn \tag{7.3.33}$$

Questa funzione, il cui andamento è visibile in Figura 7.12, viene calcolata con metodi numerici ed è generalmente disponibile nei linguaggi o ambienti di programmazione dei calcolatori elettronici. Noti i valori della funzione $\Phi(x)$, è possibile calcolare quelli della funzione distribuzione $F_X(x)$ relativa a una variabile aleatoria $X \in \mathcal{N}(\eta_X, \sigma_X^2)$ mediante la relazione (ricavabile immediatamente dalla (7.3.32))

$$F_X(x) = \Phi\left(\frac{x-\eta_X}{\sigma_X}\right). \tag{7.3.34}$$

La probabilità che una variabile aleatoria Gaussiana assuma valori in un intervallo $[a,b]$ è quindi pari a

$$\Pr\{a < X \le b\} = F_X(b) - F_X(a) = \Phi\left(\frac{b-\eta_X}{\sigma_X}\right) - \Phi\left(\frac{a-\eta_X}{\sigma_X}\right) \tag{7.3.35}$$

In pratica, quando si deve calcolare numericamente la probabilità (7.3.35) bisogna ricorrere a una valutazione *numerica* della funzione Φ. Però, sia nei programmi moderni di matematica assistita dal calcolatore (MatLab, Mathematica ecc.) sia nei linguaggi generali di programmazione ad alto livello (Fortran, C++ ecc.) è più comune avere a disposizione come funzione di macchina la cosiddetta *funzione errore* (o *error function*) $\mathrm{erf}(x)$ e la *funzione errore complementare* (o *complementary error function*) $\mathrm{erfc}(x)$, definite come segue:

$$\mathrm{erf}(x) \stackrel{\Delta}{=} \frac{2}{\sqrt{\pi}} \int_0^x e^{-\theta^2} d\theta \tag{7.3.36}$$

$$\mathrm{erfc}(x) \stackrel{\Delta}{=} 1 - \mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-\theta^2} d\theta \tag{7.3.37}$$

La Figura 7.13 mostra l'andamento delle due funzioni $\mathrm{erf}(x)$ ed $\mathrm{erfc}(x)$; si noti la proprietà di simmetria dispari della $\mathrm{erf}(x)$ attorno al punto $x = 0$. La relazione fra $\Phi(x)$ ed $\mathrm{erf}(x)$ è inoltre

$$\Phi(x) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right) \tag{7.3.38}$$

e la probabilità di un intervallo è

$$\begin{aligned}\Pr\{a < X \le b\} &= \frac{1}{2}\left[\mathrm{erf}\left(\frac{b-\eta_X}{\sqrt{2\sigma_X^2}}\right) - \mathrm{erf}\left(\frac{a-\eta_X}{\sqrt{2\sigma_X^2}}\right)\right] \\ &= \frac{1}{2}\left[\mathrm{erfc}\left(\frac{a-\eta_X}{\sqrt{2\sigma_X^2}}\right) - \mathrm{erfc}\left(\frac{b-\eta_X}{\sqrt{2\sigma_X^2}}\right)\right]\end{aligned} \tag{7.3.39}$$

CENTRO BIBLIOTECA INGEGNERIA PISA

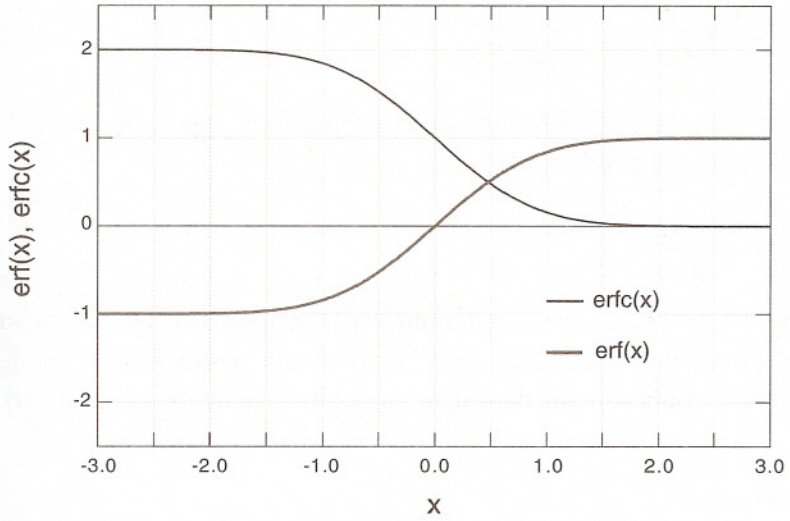


Figura 7.13 Andamento delle funzioni erf(x) ed erfc(x)

Nei problemi che coinvolgono le variabili aleatorie Gaussiane è usata talvolta anche la funzione

$$Q(x) \stackrel{\Delta}{=} 1 - \Phi(x) = \frac{1}{2}\operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) \tag{7.3.40}$$

il cui andamento è pure riportato in Figura 7.12 per completezza.

7.3.6 Variabili aleatorie condizionate

Ricordiamo che la definizione della funzione distribuzione di probabilità per una variabile aleatoria X passa attraverso la definizione di un *evento* $A_x = \{X \le x\}$ di cui in pratica la funzione rappresenta la probabilità:

$$\Pr(A) = \Pr\{X \le x\} = F_X(x) \tag{7.3.41}$$

Il verificarsi di questo evento può però in generale essere influenzato dell'essersi verificato un secondo evento B (avente probabilità non nulla). Ha senso quindi porsi il problema della caratterizzazione statistica della variabile aleatoria X *condizionatamente* al verificarsi dell'evento B considerato. Questa caratterizzazione si effettua definendo la *funzione distribuzione condizionata* $F_{X|B}(x|B)$ della variabile aleatoria X, ovvero la probabilità dell'evento A condizionato al

verificarsi di B[3]:

$$F_{X|B}(x|B) \triangleq \frac{\Pr\{A_x\, B\}}{\Pr(B)} = \frac{\Pr\{X \le x, B\}}{\Pr(B)} \tag{7.3.42}$$

Si definisce, poi, la funzione *densità di probabilità condizionata* $f_{X|B}(x|B)$ mediante la conseguente relazione

$$f_{X|B}(x|B) \triangleq \frac{dF_X(x|B)}{dx} \tag{7.3.43}$$

La funzione $F_{X|B}(x|B)$, essendo pur sempre una funzione distribuzione, gode di tutte le proprietà delle funzioni distribuzione non condizionate; analoghe considerazioni si applicano alla densità di probabilità condizionata $f_{X|B}(x|B)$.

Esempio 7.6

Risolviamo il problema del *tempo di guasto dopo rodaggio*. Supponiamo di effettuare il collaudo di una partita di lampadine, accendendole e lasciandole accese fino allo spegnimento per guasto (bruciatura del filamento). In prima approssimazione, il *tempo di guasto* così definito è una variabile aleatoria X esponenziale unilatera, avente densità di probabilità

$$f_X(x) = \frac{1}{\eta}\exp\left(-\frac{x}{\eta}\right)\mathrm{u}(x) \tag{E7.6.1}$$

ove η è il *tempo medio di guasto* (MTBF, *Mean Time Before Failure*).

Alteriamo adesso la modalità di collaudo come segue: attendiamo un tempo fisso x_0 dall'accensione, scartiamo le lampadine che a tale istante risultano già guaste, e ripetiamo il collaudo come in precedenza *sulle sole lampadine che a* x_0 *risultano funzionanti*. Questa operazione di *rodaggio* ha influenza sulla densità di probabilità del tempo di guasto della lampadine rimaste. Ciò che vogliamo determinare è la densità di probabilità *condizionata* $f_{X|B}(x|B)$, ove $B \equiv \{X \ge x_0\}$. A tal fine si deve calcolare dapprima la funzione distribuzione condizionata $F_{X|B}(x|B)$ e quindi (si veda la (7.3.42)), le probabilità $\Pr(B)$ e $\Pr\{X \le x, B\}$.

La probabilità $\Pr(B)$ è data da

[3] La notazione abbreviata $\{X \le x, B\}$ significa in realtà $\{X \le x\} \cap B$, e sarà usata per semplicità anche in seguito.

$$P(B) = \Pr\{X \geq x_0\} = 1 - \Pr\{X \leq x_0\} = 1 - F_X(x_0) \tag{E7.6.2}$$

ove $F_X(x)$ è la funzione distribuzione di X. La probabilità congiunta $\Pr\{X \leq x, B\}$, invece, è espressa da (si veda la Figura 7.14):

$$\Pr\{X \leq x, B\} = \Pr\{X \leq x, X \geq x_0\} = \begin{cases} F_X(x) - F_X(x_0) & x \geq x_0 \\ 0 & \textit{altrimenti} \end{cases} \tag{E7.6.3}$$

ovvero

$$\Pr\{X \leq x, B\} = \left[F_X(x) - F_X(x_0)\right] \mathrm{u}(x - x_0) \tag{E7.6.4}$$

Sostituendo, si ottiene

$$F_{X|B}(x|B) = \frac{\Pr\{X \leq x, B\}}{\Pr(B)} = \frac{F_X(x) - F_X(x_0)}{1 - F_X(x_0)} \mathrm{u}(x - x_0) \tag{E7.6.5}$$

e anche

$$f_{X|B}(x|B) = \frac{dF_{X|B}(x|B)}{dx} = \frac{f_X(x)}{1 - F_X(x_0)} \mathrm{u}(x - x_0) \tag{E7.6.6}$$

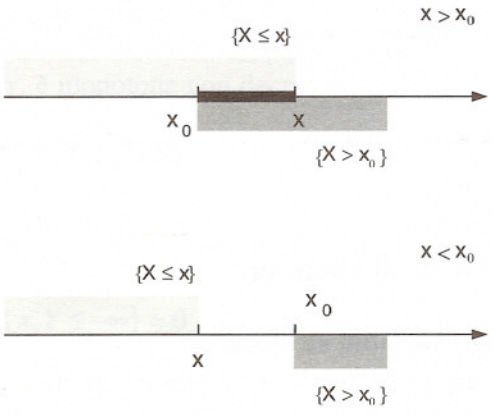


Figura 7.14 Probabilità di eventi

Questa formula spiega quali sono gli effetti del *rodaggio*: i) la densità del tempo di guasto delle lampadine sopravvissute è ovviamente nulla per $x < x_0$ (il rodaggio è infatti stato superato con certezza, e il tempo di guasto è necessariamente

maggiore di x_0); ii) la medesima densità condizionata ha lo stesso andamento della densità incondizionata per $x \geq x_0$, con l'aggiunta del fattore di scala $\left(1 - F_X(x_0)\right)^{-1}$ per ri-normalizzare a 1 la probabilità totale di guasto.

Nel nostro caso particolare di tempo di guasto esponenziale si trova

$$f_{X|B}(x|B) = \frac{1}{\eta} \exp\left(-\frac{x - x_0}{\eta}\right) \mathrm{u}(x - x_0) \tag{E7.6.7}$$

che è paragonata con la densità incondizionata nella Figura 7.15. Il lettore spieghi perché, dal punto di vista del fabbricante, non c'è alcuna convenienza di effettuare il rodaggio in fabbrica per il caso particolare delle lampadine esponenziali.

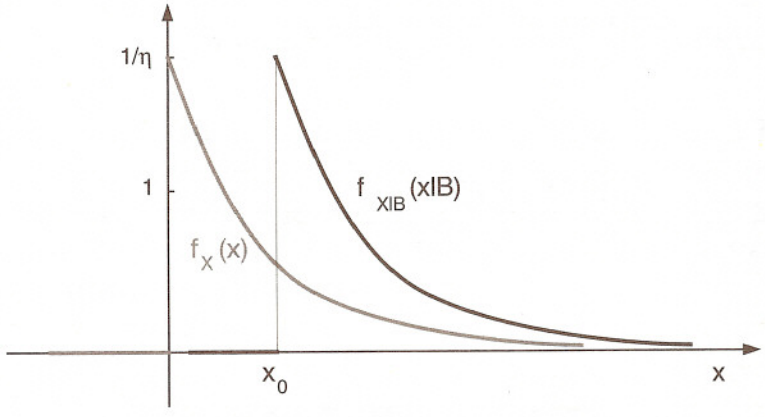


Figura 7.15 Densità di probabilità di una lampadina prima e dopo il rodaggio

❑

7.4 Sistemi di variabili aleatorie

7.4.1 Sistemi di due variabili aleatorie

Nello studio di un esperimento aleatorio può essere utile associare una *coppia* (x, y) di numeri reali ai risultati dell'esperimento stesso, definendo così un *coppia* (X, Y) di variabili aleatorie. La caratterizzazione delle due variabili aleatorie considerate *singolarmente* si può effettuare come discusso nel paragrafo precedente attraverso le funzioni distribuzione (densità) di probabilità $F_X(x)$ e $F_Y(y)$ ($f_X(x)$ e $f_Y(y)$). Queste funzioni però non danno alcuna informazione sul comportamento *congiunto* delle due variabili aleatorie. Consideriamo infatti come

variabili aleatorie il *peso* e l'*altezza* di una persona scelta casualmente in una certa popolazione. È chiaro che sarà molto improbabile trovare una persona molto alta e *congiuntamente* molto leggera. C'è un'influenza reciproca tra i valori assunti dalle due variabili che non può ovviamente essere descritta dalle sole funzioni $F_X(x)$ e $F_Y(y)$ che riguardano il *solo* peso o la *sola* altezza senza minimamente tener conto dell'altra grandezza. È importante, allora, disporre di una caratterizzazione statistica *congiunta* di tali variabili.

Data la coppia (X,Y) di variabili aleatorie, si definisce la *funzione distribuzione di probabilità congiunta*

$$F_{XY}(x,y) \triangleq \Pr\{X \le x, Y \le y\} \tag{7.4.1}$$

la quale descrive in modo completo il comportamento statistico congiunto delle due variabili. La funzione $F_{XY}(x,y)$, come vedremo, determina anche le proprietà statistiche *marginali*, cioè relative a una sola variabile della coppia.

Elenchiamo, adesso, alcune proprietà importanti della funzione $F_{XY}(x,y)$:

- la funzione $F_{XY}(x,y)$ assume valori compresi tra 0 e 1, ovvero

$$0 \le F_{XY}(x,y) \le 1 \tag{7.4.2}$$

- la funzione $F_{XY}(x,y_0)$, comunque si scelga il valore y_0 della variabile y, è monotona non decrescente nella variabile x e continua da destra in questa variabile; analogamente, la funzione $F_{XY}(x_0,y)$, comunque si scelga il valore x_0 della variabile x, è monotona non decrescente e continua da destra nella variabile y;
- la funzione $F_{XY}(x,y)$ soddisfa le uguaglianze

$$F_{XY}(-\infty,y) = P\{X \le -\infty, Y \le y\} = 0 \tag{7.4.3a}$$

$$F_{XY}(x,-\infty) = P\{X \le x, Y \le -\infty\} = 0 \tag{7.4.3b}$$

e naturalmente anche

$$F_{XY}(-\infty,-\infty) = 0 \tag{7.4.3c}$$

- le funzioni distribuzione *marginali* delle variabili aleatorie X e Y si ricavano dalla congiunta come segue:

$$F_Y(y) = F_{XY}(+\infty,y) \tag{7.4.4a}$$

$$F_X(x) = F_{XY}(x, +\infty) \tag{7.4.4b}$$

- il limite della funzione $F_{XY}(x,y)$ quando sia x sia $y \to +\infty$ è unitario, cioè

$$F_{XY}(+\infty, +\infty) = 1 \tag{7.4.5}$$

- la probabilità dell'evento *rettangolare* $R = \{x_1 < X \le x_2, y_1 < Y \le y_2\}$ può essere calcolata mediante la relazione

$$\begin{aligned} &\Pr\{x_1 < X \le x_2\ y_1 < Y \le y_2\} \\ &= F_{XY}(x_2, y_2) - F_{XY}(x_1, y_2) - F_{XY}(x_2, y_1) + F_{XY}(x_1, y_1) \end{aligned} \tag{7.4.6}$$

Riprendiamo quest'ultima relazione considerando un evento rettangolare di misura molto piccola, avente cioè "lati" di ampiezza rispettivamente Δx e Δy prossime a zero. La probabilità di questo evento è dunque

$$\begin{aligned} &\Pr\{x < X \le x + \Delta x, y < Y \le y + \Delta y\} \\ &= F_{XY}(x + \Delta x, y + \Delta y) - F_{XY}(x, y + \Delta y) - \left[F_{XY}(x + \Delta x, y) - F_{XY}(x, y)\right] \\ &\cong \frac{\partial F_{XY}(x, y + \Delta y)}{\partial x}\Delta x - \frac{\partial F_{XY}(x, y)}{\partial x}\Delta x = \frac{\partial^2 F_{XY}(x, y)}{\partial x\, \partial y}\Delta x \Delta y \end{aligned} \tag{7.4.7}$$

ove naturalmente l'approssimazione è valida nella misura in cui Δx e Δy sono "piccoli". Se definiamo la funzione

$$f_{XY}(x,y) \triangleq \frac{\partial^2 F_{XY}(x,y)}{\partial x \partial y} \tag{7.4.8}$$

abbiamo allora

$$\Pr\{x < X \le x + \Delta x, y < Y \le y + \Delta y\} \cong f_{XY}(x,y)\Delta x \Delta y \tag{7.4.9}$$

o anche

$$f_{XY}(x,y) \cong \frac{\Pr\{x < X \le x + \Delta x, y < Y \le y + \Delta y\}}{\Delta x \Delta y} \tag{7.4.10}$$

Quest'ultima relazione giustifica pienamente il nome di *funzione densità di probabilità congiunta* che si dà alla $f_{XY}(x,y)$ definita come nella (7.4.8). Si noti l'analogia con la definizione (7.3.10) della densità *marginale* $f_X(x)$.

La funzione densità di probabilità congiunta $f_{XY}(x,y)$ gode delle proprietà seguenti:

- essa assume valori non negativi, ovvero

$$f_{XY}(x,y) \geq 0 \tag{7.4.11}$$

- essa integra a 1 su tutto il piano:

$$\int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{+\infty} f_{XY}(x,y)\, dx\, dy = 1 \tag{7.4.12}$$

- le densità marginali $f_X(x)$ ed $f_Y(y)$ rispettivamente delle variabili aleatorie X e Y si possono ricavare come segue:

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dy \tag{7.4.13a}$$

$$f_Y(y) = \int_{-\infty}^{+\infty} f_{XY}(x,y)\, dx \tag{7.4.13b}$$

- la probabilità di un evento $A = \{(X,Y) \in D\}$ individuato da un dominio D nel piano cartesiano (x, y) è data da

$$\Pr(A) = \iint_D f_{XY}(x,y)\, dx\, dy \tag{7.4.14}$$

- la funzione distribuzione congiunta $F_{XY}(x,y)$ può essere ricavata dalla funzione densità di probabilità congiunta $f_{XY}(x,y)$ mediante la relazione

$$F_{XY}(x,y) = \int_{\alpha=-\infty}^{x} \int_{\beta=-\infty}^{y} f_{XY}(\alpha,\beta)\, d\alpha\, d\beta \tag{7.4.15}$$

7.4.2 Funzioni distribuzione e densità di probabilità condizionate

Consideriamo una coppia di variabili aleatorie (X,Y) con densità di probabilità congiunta $f_{XY}(x,y)$ e supponiamo di aver osservato che la variabile X ha assunto il particolare valore x; la distribuzione marginale della variabile Y viene modificata da questo *condizionamento*. Definiamo allora la *funzione distribuzione condizionata* della variabile aleatoria Y rispetto all'evento $\{X = x\}$ come segue:

$$F_{Y|X}(y|x) \triangleq \frac{\int_{\beta=-\infty}^{y} f_{XY}(x,\beta)\, d\beta}{f_X(x)} \tag{7.4.16}$$

La funzione *densità di probabilità condizionata* $f_{Y|X}(y|x)$ della variabile aleatoria Y, rispetto all'evento $\{X = x\}$, si ricava poi derivando la funzione

distribuzione condizionata $F_{Y|X}(y|x)$ rispetto alla variabile y:

$$f_{Y|X}(y|x) \triangleq \frac{dF_Y(y|x)}{dy} = \frac{f_{XY}(x,y)}{f_X(x)} \tag{7.4.17}$$

Se la densità di probabilità *marginale* $f_Y(y)$ della variabile aleatoria X coincide con la densità di probabilità *condizionata* $f_{Y|X}(y|x)$ il comportamento statistico della variabile Y non è influenzato dal valore assunto dalla variabile X. In tal caso le variabili aleatorie X e Y sono *indipendenti*. Se $f_{Y|X}(y|x) = f_Y(y)$, allora la densità di probabilità congiunta $f_{XY}(x,y)$ può essere espressa come prodotto delle densità di probabilità marginali $f_Y(y)$ ed $f_X(x)$, cioè

$$f_{XY}(x,y) = f_X(x) f_Y(y) \tag{7.4.18}$$

7.4.3 Trasformazione di una coppia di variabili aleatorie

Definiamo, ora, una variabile aleatoria Z come funzione di una coppia di variabili (X,Y), aventi densità di probabilità congiunta $f_{XY}(x,y)$:

$$Z = g(X,Y) \tag{7.4.19}$$

ove $g(\cdot,\cdot)$ è una funzione reale di due variabili reali. La funzione distribuzione $F_Z(z)$ della variabile aleatoria Z è

$$F_Z(z) = \Pr\{Z \le z\} = \Pr\{g(X,Y) \le z\} \tag{7.4.20}$$

Il calcolo della funzione $F_Z(z)$ può essere effettuato, allora, con la relazione (7.4.14), cioè

$$F_Z(z) = \iint_{R(z)} f_{XY}(x,y)\, dx\, dy \tag{7.4.21}$$

ove $R(z)$ indica la regione del piano cartesiano individuata dall'evento $\{g(X,Y) \le z\}$, ovvero l'insieme dei punti (x,y) verificanti la disuguaglianza $g(x,y) \le z$. Nota la funzione $F_Z(z)$ si può ricavare la densità di probabilità $f_Z(z)$ per derivazione:

$$f_Z(z) = \frac{dF_Z(z)}{dz} \tag{7.4.22}$$

Ad esempio, se definiamo la variabile Z come *somma* di X e Y si ha (Figura 7.16)

$$F_Z(z) = \iint_{x+y\le z} f_{XY}(x,y)\,dx\,dy = \int_{x=-\infty}^{+\infty}\int_{y=-\infty}^{z-x} f(x,y)\,dx\,dy \tag{7.4.23}$$

e quindi

$$f_Z(z) = \frac{dF_Z(z)}{dz} = \int_{x=-\infty}^{+\infty} f_{XY}(x, z-x)\,dx \tag{7.4.24}$$

Se infine le due variabili aleatorie X e Y sono *indipendenti*, la densità di probabilità congiunta $f_{XY}(x,y)$ può essere fattorizzata e si ottiene

$$f_Z(z) = \int_{x=-\infty}^{+\infty} f_{XY}(x, z-x)\,dx\,dy = \int_{x=-\infty}^{+\infty} f_X(x) f_Y(z-x)\,dx = f_X(z) \otimes f_Y(z) \tag{7.4.25}$$

cioè la densità di $Z = X + Y$ è pari alla *convoluzione* delle due densità marginali delle variabili X e Y.

Tornando al caso generale, notiamo infine che per il calcolo del valore atteso η_Z della variabile aleatoria Z si può utilizzare la formula

$$\eta_Z = \mathrm{E}\{Z\} = \mathrm{E}\{g(X,Y)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{XY}(x,y)\,dx\,dy \tag{7.4.26}$$

che rappresenta una generalizzazione del teorema del valor medio (7.3.24).

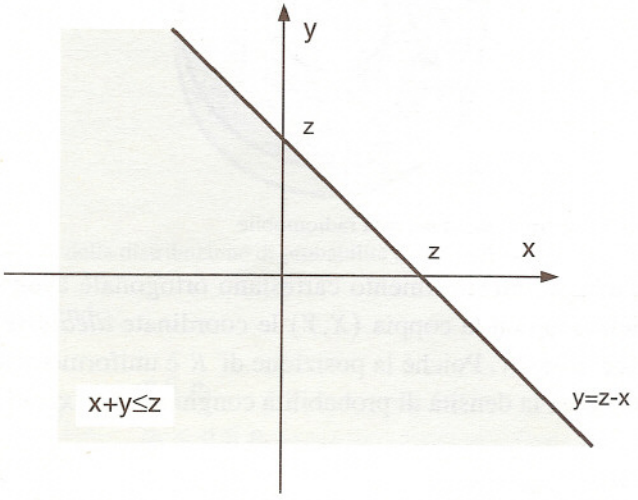


Figura 7.16 Calcolo della distribuzione della somma di due variabili aleatorie

Esempio 7.7

La Figura 7.17 rappresenta la generica "cella" a corona circolare di un sistema radio cellulare. La stazione base, posta nel punto S, trasmette al generico ricevitore R situato a una distanza aleatoria D (maggiore di r_1 e minore di r_2). Il ricevitore riceve un segnale di potenza inversamente proporzionale al quadrato della distanza dal trasmettitore (con costante di proporzionalità k nota). Supponendo che la posizione del ricevitore sia uniformemente distribuita all'interno della cella, determiniamo la funzione densità di probabilità $f_P(p)$ della potenza P ricevuta.

La potenza P di segnale ricevuta da R è espressa dalla relazione

$$P = \frac{k}{D^2} \tag{E7.7.1}$$

La funzione distribuzione della variabile aleatoria P è data, allora, da

$$F_P(p) = \Pr\{P \le p\} = \Pr\left\{\frac{k}{D^2} \le p\right\} = \Pr\left\{D \ge \sqrt{\frac{k}{p}}\right\} \tag{E7.7.2}$$

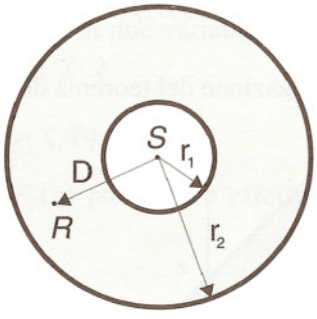


Figura 7.17 Cella di copertura di un sistema radiomobile

Consideriamo, adesso, un riferimento cartesiano ortogonale avente origine nel punto S e indichiamo con la coppia (X,Y) le coordinate *aleatorie* del punto in cui si trova il ricevitore R. Poiché la posizione di R è uniformemente distribuita all'interno della cella, la densità di probabilità congiunta $f_{XY}(x,y)$ è pari a

$$f_{XY}(x,y) = \frac{1}{\pi\left(r_2^2 - r_1^2\right)} \tag{E7.7.3}$$

per tutti i punti interni alla corona circolare, e

$$f_{XY}(x,y)=0 \tag{E7.7.4}$$

al di fuori di essa. La funzione distribuzione di P è allora

$$F_P(p)=\Pr\left\{D\geq\sqrt{\frac{k}{p}}\right\}=\Pr\left\{\sqrt{X^2+Y^2}\geq\sqrt{\frac{k}{p}}\right\}=\int_{A(p)} f_{XY}(x,y)\,dx\,dy \tag{E7.7.5}$$

ove $A(p)$ è il dominio

$$A(p)=\left\{(x,y)\left|\sqrt{x^2+y^2}\geq\sqrt{\frac{k}{p}}\right.\right\} \tag{E7.7.6}$$

ovvero è l'insieme dei punti del piano la cui distanza dall'origine non è inferiore a $\sqrt{k/p}$ (si veda la Figura 7.18). Per comodità, identifichiamo le potenze ricevute ai bordi della cella: $p_1 \triangleq k/r_1^2$ e $p_2 \triangleq k/r_2^2$. Svolgendo l'integrale sulla corona circolare di Figura 7.18, si ricava che, quando $p_2 < p \leq p_1$,

$$\int_{A(p)} f_{XY}(x,y)\,dx\,dy=\frac{\pi r_2^2-\pi k/p}{\pi\left(r_2^2-r_1^2\right)}=\frac{1-p_2/p}{1-p_2/p_1} \tag{E7.7.7}$$

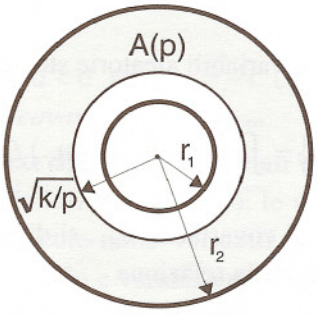


Figura 7.18 Calcolo della distribuzione di probabilità dell'Esempio 7.7

e quindi, riassumendo,

$$F_P(p)=\begin{cases}0 & p\leq p_2\\ \dfrac{1-p_2/p}{1-p_2/p_1} & p_2<p\leq p_1\\ 1 & p>p_1\end{cases} \tag{E7.7.8}$$

Derivando la funzione distribuzione $F_P(p)$ si trova infine la seguente

espressione della densità di probabilità $f_P(p)$:

$$f_P(p) = \begin{cases} \dfrac{p_2}{1 - p_2/p_1} \cdot \dfrac{1}{p^2} & p_2 < p \le p_1 \\ 0 & \text{altrimenti} \end{cases} \tag{E7.7.9}$$

❑

7.4.4 Correlazione e covarianza

Come abbiamo discusso nelle pagine precedenti, il comportamento statistico di una variabile aleatoria X può essere caratterizzato in maniera incompleta ma talvolta sufficiente da alcuni parametri caratteristici, quali il valore atteso η_X e la varianza σ_X^2. Analogamente, per una *coppia* di variabili aleatorie (X,Y), è possibile determinare alcuni parametri statistici semplificati che rappresentano utili indicazioni per la comprensione del loro comportamento statistico *congiunto*. Indici molto importanti sono la *correlazione* r_{XY} tra le variabili aleatorie X e Y:

$$r_{XY} \triangleq \mathrm{E}\{XY\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x\, y\, f_{XY}(x,y)\, dx\, dy \tag{7.4.27}$$

e la *covarianza* c_{XY} tra le due variabili aleatorie stesse:

$$c_{XY} \triangleq \mathrm{E}\{(X - \eta_X)(Y - \eta_Y)\} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \eta_X)(y - \eta_Y)\, f_{XY}(x,y)\, dx\, dy \tag{7.4.28}$$

Sviluppando la definizione di covarianza c_{XY}, si dimostra facilmente che questa è legata alla correlazione r_{XY} dalla relazione

$$c_{XY} = r_{XY} - \eta_X \eta_Y \tag{7.4.29}$$

La covarianza è un parametro statistico molto importante che tende ad accertare se tra le due variabili X e Y esiste una relazione di dipendenza di tipo *lineare,* e che comunque misura la tendenza di variazione congiunta (*co*-varianza) delle due. Se la covarianza è grande e positiva, le due variabili aleatorie X e Y tendono a discostarsi dal rispettivo valor medio *nella stessa direzione*, cioè le due quantità $(X - \eta_X)$ e $(Y - \eta_Y)$ tendono ad avere lo stesso segno. È questo il caso, ad esempio, del *peso* e dell'*altezza* di una persona scelta a caso: se l'altezza è maggiore della media, così sarà anche presumibilmente il peso. Viceversa, cova-

rianza negativa indica versi di variazione opposti (ad esempio, *età* e *acuità visiva*). Se la covarianza tra due variabili aleatorie è *nulla*, le variabili si dicono *incorrelate*.

Il medesimo significato della covarianza ha il *coefficiente di correlazione* ρ_{XY} fra le variabili aleatorie X e Y:

$$\rho_{XY} \triangleq \mathrm{E}\left\{\frac{X-\eta_X}{\sigma_X}\cdot\frac{Y-\eta_Y}{\sigma_Y}\right\} = \frac{c_{XY}}{\sigma_X\sigma_Y} = \frac{r_{XY}-\eta_X\eta_Y}{\sigma_X\sigma_Y} \tag{7.4.30}$$

che gode delle proprietà seguenti:

- il suo modulo non può assumere valori maggiori dell'unità:

$$|\rho_{XY}| \le 1 \tag{7.4.31}$$

- esso assume valore nullo se e solo se le variabili aleatorie X e Y sono incorrelate;
- il suo modulo assume valore unitario se e solo se le variabili aleatorie X e Y (che in tal caso si dicono *completamente correlate*) sono *linearmente dipendenti*, cioè se sono legate da una relazione del tipo

$$Y = aX + b \tag{7.4.32}$$

con $a>0$ se $\rho_{XY}=1$ e $a<0$ se $\rho_{XY}=-1$ (si ricordi la discussione sul significato del parametro *covarianza*).

Osserviamo che il coefficiente di correlazione è un valore di *covarianza normalizzata,* come si vede dalla definizione (7.4.30): le variabili aleatorie X e Y vengono trasformate rispettivamente nelle variabili $(X-\eta_X)/\sigma_X$ e $(X-\eta_Y)/\sigma_Y$ entrambe aventi valor medio nullo e varianza unitaria per poter definire un parametro di correlazione *universale*, cioè omogeneo per coppie di variabili anche molto diverse come valori numerici. In conseguenza di questa operazione, ρ_{XY} è sempre limitato in ampiezza all'intervallo [0,1], e ciò rende ogni coppia di variabili "commensurabile" con ogni altra. Se il coefficiente ρ_{XY} è (in modulo) vicino a 1, le variabili aleatorie X e Y tendono a seguire una relazione *lineare* di variazione reciproca; viceversa, se $\rho_{XY}=0$, le variabili sono *incorrelate* e la tendenza reciproca dei valori delle due variabili aleatorie non è di tipo lineare.

Quando le variabili aleatorie X e Y sono *indipendenti*, la loro correlazione è

$$r_{XY} = E\{XY\} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f_{XY}(x,y)\,dx\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\, f_X(x) f_Y(y)\,dx\,dy =$$

$$= \int_{-\infty}^{+\infty} x\, f_X(x)\, dx \int_{-\infty}^{+\infty} y f_Y(y)\, dy = \eta_X\, \eta_Y \tag{7.4.33}$$

Si trova quindi $c_{XY} = r_{XY} - \eta_X \eta_Y = 0$, ovvero *due variabili aleatorie indipendenti sono anche incorrelate*. L'implicazione inversa, tuttavia, non è vera. Il lettore provi a dimostrare come esempio che le due variabili $X \in \mathcal{U}[-1,1]$ e $Y = X^2$, pur non essendo indipendenti, sono *incorrelate*. In altre parole, l'indipendenza è una condizione *più restrittiva* dell'incorrelazione.

7.4.5 Sistemi di *n* variabili aleatorie e vettori aleatori

Alcuni concetti e risultati relativi a una coppia di variabili aleatorie possono essere facilmente estesi al caso di un *sistema* $(X_1, X_2, ..., X_n)$ costituito da n variabili aleatorie, cioè al caso di una variabile aleatoria n-dimensionale. In perfetta analogia al caso bidimensionale già esaminato in dettaglio, la *funzione distribuzione di probabilità congiunta* di tale sistema è definita come segue:

$$F_{X_1 X_2 ... X_n}(x_1, x_2, ..., x_n) \triangleq \Pr\{X_1 \le x_1, X_2 \le x_2, ..., X_n \le x_n\} \tag{7.4.34}$$

e la relativa *funzione densità di probabilità congiunta* è

$$f_{X_1 X_2 ... X_n}(x_1, x_2, ..., x_n) \triangleq \frac{\partial^n F_{X_1 X_2 ... X_n}(x_1, x_2, ..., x_n)}{\partial x_1 \partial x_2 ... \partial x_n} \tag{7.4.35}$$

Data la densità di probabilità congiunta $f_{X_1 X_2 ... X_n}(x_1, x_2, ..., x_n)$, è possibile ricavare la densità marginale di ciascuna variabile o le densità congiunte di un sottoinsieme del sistema mediante relazioni simili a quelle illustrate per la coppia di variabili aleatorie. Per ricavare, ad esempio, la funzione densità di probabilità congiunta del (sotto-)sistema $(X_1, X_3, ..., X_n)$ basta integrare la densità congiunta (7.4.35) rispetto alla variabile mancante nel sottogruppo:

$$f_{X_1 X_3 ... X_n}(x_1, x_3, ..., x_n) = \int_{-\infty}^{+\infty} f_{X_1 X_2 X_3 ... X_n}(x_1, x_2, x_3 ..., x_n)\, dx_2 \tag{7.4.36}$$

Considerazioni analoghe possono essere ripetute per le *funzioni densità di probabilità condizionate*. Per calcolare, ad esempio, la densità di probabilità congiunta delle variabili aleatorie $X_1, X_4, ..., X_n$ condizionata rispetto alle variabili aleatorie X_2, X_3 si utilizza la relazione

$$f_{X_1 X_4 ... X_n | X_2 X_3}(x_1, x_4, ..., x_n | x_2, x_3) = \frac{f_{X_1 X_2 X_3 X_4 ... X_n}(x_1, x_2, x_3, ..., x_n)}{f_{X_2 X_3}(x_2, x_3)} \tag{7.4.37}$$

Le variabili aleatorie sono infine *indipendenti* se la densità di un *qualunque* sottogruppo di esse condizionata a un *qualunque* altro sottogruppo (ovviamente costituito da variabili aleatorie distinte dalle prime) è pari alla relativa densità *incondizionata*.

Nello studio dei sistemi di variabili aleatorie n-dimensionali si utilizza, di solito, una notazione più compatta. Le n variabili aleatorie $(X_1, X_2, ...X_n)$, infatti, vengono disposte in un *vettore aleatorio* $\mathbf{X}$:

$$\mathbf{X} \triangleq \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = [X_1, X_2, \ldots, X_n]^T \tag{7.4.38}$$

ove $[\cdot]^T$ indica l'operatore di trasposizione. Le funzioni distribuzione (7.4.34) e densità di probabilità (7.4.35) possono essere allora indicate rispettivamente con la scrittura $F_{\mathbf{X}}(\mathbf{x})$ e $f_{\mathbf{X}}(\mathbf{x})$. Per semplificare la notazione, anche gli *indici caratteristici* delle variabili possono essere rappresentati con una notazione vettoriale. Il vettore valor medio $\boldsymbol{\eta}_{\mathbf{X}}$ del vettore $\mathbf{X}$ (7.4.38), ad esempio, è definito dalla relazione

$$\boldsymbol{\eta}_{\mathbf{X}} \triangleq \mathrm{E}\{\mathbf{X}\} = [\eta_{X_1}, \eta_{X_2}, \ldots, \eta_{X_n}]^T \tag{7.4.39}$$

ed è quindi pari al vettore colonna dei valori attesi delle n variabili. Come nel caso bidimensionale, utili informazioni sul comportamento statistico del vettore aleatorio possono essere acquisite attraverso la conoscenza del valore della correlazione e/o della covarianza di *tutte le coppie* di variabili aleatorie estraibili dal vettore $\mathbf{X}$. Le correlazioni $\{r_{X_i X_j}; i, j = 1, \ldots, n\}$ fra tutte le variabili aleatorie di $\mathbf{X}$ possono essere raccolte in una matrice di dimensioni $n \times n$ $\mathbf{R}_{\mathbf{X}}$, detta *matrice di correlazione,* e definita da

$$\mathbf{R}_{\mathbf{X}} \triangleq \mathrm{E}\{\mathbf{X}\mathbf{X}^T\} = \begin{bmatrix} r_{X_1X_1} & r_{X_1X_{21}} & \cdots & r_{X_1X_n} \\ r_{X_2X_1} & r_{X_2X_2} & \cdots & r_{X_2X_n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{X_nX_1} & r_{X_nX_2} & \cdots & r_{X_nX_n} \end{bmatrix} \tag{7.4.40}$$

Tale matrice è simmetrica, essendo $\mathrm{E}\{X_i X_j\} = \mathrm{E}\{X_j X_i\}$, e gli elementi disposti lungo la sua diagonale principale sono i valori quadratici medi delle variabili aleatorie costituenti il sistema, in quanto $r_{X_iX_i} = \mathrm{E}\{X_i X_i\} = m_{X_i}^2$. Si definisce,

analogamente, la *matrice di covarianza* $\mathbf{C}_\mathbf{X}$ del vettore $\mathbf{X}$:

$$\mathbf{C}_\mathbf{X} \triangleq \mathrm{E}\left\{(\mathbf{X}-\boldsymbol{\eta}_\mathbf{X})(\mathbf{X}-\boldsymbol{\eta}_\mathbf{X})^T\right\} = \begin{bmatrix} c_{X_1X_1} & c_{X_1X_2} & \cdots & c_{X_1X_n} \\ c_{X_2X_1} & c_{X_2X_2} & \cdots & c_{X_2X_n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{X_nX_1} & c_{X_nX_2} & \cdots & c_{X_nX_n} \end{bmatrix} \tag{7.4.41}$$

Anche la matrice di covarianza è simmetrica, e l'i-esimo elemento della sua diagonale principale rappresenta la varianza $\sigma^2_{X_i}$ della variabile aleatoria X_i, essendo $c_{X_iX_i} = \mathrm{E}\{(X_i - \eta_{X_i})^2\}$. La matrice di covarianza (7.4.41), si può quindi riscrivere come

$$\mathbf{C}_\mathbf{X} = \begin{bmatrix} \sigma^2_{X_1} & c_{X_1X_2} & \cdots & c_{X_1X_n} \\ c_{X_2X_1} & \sigma^2_{X_2} & \cdots & c_{X_2X_n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{X_nX_1} & c_{X_nX_2} & \cdots & \sigma^2_{X_n} \end{bmatrix} = \mathbf{R}_\mathbf{X} - \boldsymbol{\eta}_\mathbf{X}\boldsymbol{\eta}_\mathbf{X}^T \tag{7.4.42}$$

7.4.6 Trasformazione di un vettore aleatorio

Consideriamo un vettore aleatorio $\mathbf{Y}$ n-dimensionale espresso come funzione di un altro vettore aleatorio $\mathbf{X}$ di ugual dimensione:

$$\mathbf{Y} = \mathbf{g}(\mathbf{X}) \tag{7.4.43}$$

ove $\mathbf{g}(\bullet)$ è una funzione reale n-dimensionale di n variabili reali. Per determinare la funzione densità di probabilità congiunta $f_\mathbf{Y}(\mathbf{y})$ del vettore $\mathbf{Y}$, nota la densità congiunta $f_\mathbf{X}(\mathbf{x})$ del vettore $\mathbf{X}$, si può utilizzare il cosiddetto *teorema fondamentale generalizzato*:

$$f_\mathbf{Y}(\mathbf{y}) = \sum_i \frac{f_\mathbf{X}(\mathbf{x}_i)}{\left|\det\left(\mathbf{J}(\mathbf{x}_i)\right)\right|} \tag{7.4.44}$$

ove $\{\mathbf{x}_i\}$ è l'insieme di tutte le possibili soluzioni, per un vettore $\mathbf{y}$ assegnato, del sistema di equazioni

$$\mathbf{g}(\mathbf{x}_i) = \mathbf{y} \tag{7.4.45}$$

e dove la quantità $\mathbf{J}(\mathbf{x}_i)$ rappresenta la *matrice Jacobiana* della trasformazione

(7.4.43) calcolata per $\mathbf{x} = \mathbf{x}_i$:

$$\mathbf{J}(\mathbf{x}_i) \triangleq \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{bmatrix}_{\mathbf{x}=\mathbf{x}_i} \tag{7.4.46}$$

Purtroppo il teorema fondamentale è applicabile nella forma appena vista solo quando la dimensione n del vettore aleatorio trasformato $\mathbf{Y}$ è uguale a quella del vettore di partenza $\mathbf{X}$. Un altro caso importante è però quello in cui il vettore $\mathbf{X}$ viene trasformato in un'*unica* variabile aleatoria monodimensionale Z:

$$Z = g(\mathbf{X}) = g(X_1, X_2, \ldots, X_n) \tag{7.4.47}$$

ove $g(\cdot)$ è una funzione reale monodimensionale di n variabili reali. La densità di probabilità $f_Z(z)$ della variabile aleatoria Z può essere ricavata con il metodo illustrato nel Paragrafo 7.4.3. Si calcola, innanzitutto, la funzione distribuzione $F_Z(z)$ (si veda la (7.4.21)) mediante la relazione

$$F_Z(z) = \int_{R(z)} f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x} \tag{7.4.48}$$

ove $R(z)$ indica la regione dello spazio n-dimensionale individuata dall'evento $\{g(X_1, X_2, \ldots, X_n) \leq z\}$. Nota la funzione $F_Z(z)$, è possibile poi determinare la densità di probabilità $f_Z(z)$ della variabile Z per derivazione.

Esempio 7.8

Troviamo valore atteso e varianza della somma di n variabili aleatorie X_i, $i = 1, \ldots, n$. Conviene pensare a questo problema come la trasformazione di un vettore aleatorio $\mathbf{X}$ in una singola variabile aleatoria Z:

$$Z = \sum_{i=1}^{n} X_i \tag{E7.8.1}$$

Con notazione vettoriale, possiamo scrivere:

$$Z = \mathbf{1}^T \mathbf{X} \quad , \quad \mathbf{1} \triangleq \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \tag{E7.8.2}$$

Calcoliamo ora il valore atteso η_Z:

$$\eta_Z = \mathrm{E}\{Z\} = \mathrm{E}\{\mathbf{1}^T\mathbf{X}\} = \mathbf{1}^T\,\mathrm{E}\{\mathbf{X}\} = \mathbf{1}^T\eta_{\mathbf{X}} = \sum_{i=1}^{n}\eta_{X_i} \tag{E7.8.3}$$

che risulta dato dalla somma dei valori attesi delle n variabili X_i di cui facciamo la somma. Per quanto riguarda la *varianza* σ_Z^2 abbiamo:

$$\begin{aligned}\sigma_Z^2 &= \mathrm{E}\left\{(Z-\eta_Z)^2\right\} = \mathrm{E}\left\{(\mathbf{1}^T\mathbf{X}-\mathbf{1}^T\eta_{\mathbf{X}})(\mathbf{1}^T\mathbf{X}-\mathbf{1}^T\eta_{\mathbf{X}})^T\right\} \\ &= \mathbf{1}^T\,\mathrm{E}\left\{(\mathbf{X}-\eta_{\mathbf{X}})(\mathbf{X}-\eta_{\mathbf{X}})^T\right\}\mathbf{1} = \mathbf{1}^T\mathbf{C}_{\mathbf{X}}\mathbf{1} = \sum_{i=1}^{n}\sum_{j=1}^{n}c_{X_iX_j}\end{aligned} \tag{E7.8.4}$$

Per ottenere la varianza di Z si devono cioè sommare tutti gli elementi della matrice di covarianza delle variabili aleatorie X_i date. Come caso particolare, se tali variabili sono a due a due *incorrelate* (oppure, a maggior ragione, *indipendenti*), la varianza della loro somma è pari alla *somma delle varianze*. ❑

Esempio 7.9

Sono date le variabili aleatorie X_1 e X_2, indipendenti ed entrambe $\in \mathcal{N}(0,\sigma^2)$, a partire dalle quali si costruiscono le nuove variabili

$$Y_1 = X_1^2 + X_2^2 \tag{E7.9.1}$$

$$Y_2 = \frac{X_1}{X_2} \tag{E7.9.2}$$

Dopo aver ricavato la densità di probabilità $f_{Y_1}(y)$ della variabile aleatoria Y_1, dimostriamo che Y_1 e Y_2 sono ancora *indipendenti*.

Per dimostrare l'indipendenza delle variabili aleatorie Y_1 e Y_2 basta verificare la validità dell'uguaglianza (si veda la (7.4.18))

$$f_{Y_1Y_2}(y_1,y_2) = f_{Y_1}(y_1)f_{Y_2}(y_2) \tag{E7.9.3}$$

ove $f_{Y_1Y_2}(y_1,y_2)$ rappresenta la densità congiunta delle variabili aleatorie Y_1 e Y_2 mentre $f_{Y_1}(y_1)$, $f_{Y_2}(y_2)$ sono le relative densità marginali. Per ricavare la densità $f_{Y_1Y_2}(y_1,y_2)$ si può utilizzare il teorema fondamentale multidimensionale (7.4.44). Si devono quindi per prima cosa ricavare i punti $\mathbf{x}_i$ risolvendo il sistema nonlineare

$$y_1 = x_1^2 + x_2^2 \tag{E7.9.4}$$

$$y_2 = \frac{x_1}{x_2} \tag{E7.9.5}$$

nelle incognite x_1 e x_2. Se $y_1 < 0$ il sistema non ammette alcuna soluzione, e quindi la densità congiunta è nulla. Viceversa, se $y_1 > 0$ il sistema (E7.9.4-5) ammette le due soluzioni seguenti:

$$\mathbf{x}_1 = (x_1, x_2)_1 = \left(y_2 \sqrt{\frac{y_1}{1+y_2^2}}, \sqrt{\frac{y_1}{1+y_2^2}} \right) \tag{E7.9.6}$$

$$\mathbf{x}_2 = (x_1, x_2)_2 = \left(-y_2 \sqrt{\frac{y_1}{1+y_2^2}}, -\sqrt{\frac{y_1}{1+y_2^2}} \right) \tag{E7.9.7}$$

Il determinante dello Jacobiano della trasformazione è dato da

$$\det\left(\mathbf{J}(x_1, x_2)\right) = \det\begin{bmatrix} 2x_1 & 2x_2 \\ 1/x_2 & -x_1/x_2^2 \end{bmatrix} = -2\left[\frac{x_1^2}{x_2^2} + 1\right] = -2\left(1 + y_2^2\right) \tag{E7.9.8}$$

e quindi

$$f_{Y_1Y_2}(y_1, y_2) = \left[\frac{f_{X_1X_2}(\mathbf{x}_1)}{2\left(1+y_2^2\right)} + \frac{f_{X_1X_2}(\mathbf{x}_2)}{2\left(1+y_2^2\right)} \right] \mathrm{u}(y_1) \tag{E7.9.9}$$

La densità congiunta $f_{X_1X_2}(x_1, x_2)$ di X_1 e X_2 si ricava immediatamente tenendo conto dell'indipendenza:

$$f_{X_1X_2}(x_1, x_2) = f_{X_1}(x_1) f_{X_2}(x_2) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x_1^2 + x_2^2}{2\sigma^2} \right] \tag{E7.9.10}$$

per cui troviamo il risultato cercato:

$$f_{Y_1Y_2}(y_1, y_2) = \frac{1}{\pi} \cdot \frac{1}{1+y_2^2} \cdot \frac{1}{2\sigma^2} e^{-\frac{y_1}{2\sigma^2}} \mathrm{u}(y_1) \tag{E7.9.11}$$

La densità (marginale) della sola Y_1 si ricava subito per integrazione:

$$f_{Y_1}(y_1) = \int_{-\infty}^{+\infty} f_{Y_1Y_2}(y_1, y_2) dy_2 = \frac{1}{2\sigma^2} e^{-\frac{y_1}{2\sigma^2}} \mathrm{u}(y_1) \tag{E7.9.12}$$

ed è quella di una variabile esponenziale. Da questo segue che è possibile esprimere la densità congiunta $f_{Y_1Y_2}(y_1, y_2)$ come il prodotto di due fattori,

ciascuno dipendente solo da una variabile:

$$f_{Y_1Y_2}(y_1, y_2) = f_{Y_1}(y_1) \cdot f_{Y_2}(y_2) \tag{E7.9.13}$$

ove la densità di Y_2 è *di Cauchy* (Figura 7.19):

$$f_{Y_2}(y_2) = \frac{1}{\pi(1+y_2^2)} \tag{E7.9.14}$$

Resta quindi dimostrato che le variabili aleatorie Y_1 e Y_2 sono indipendenti.

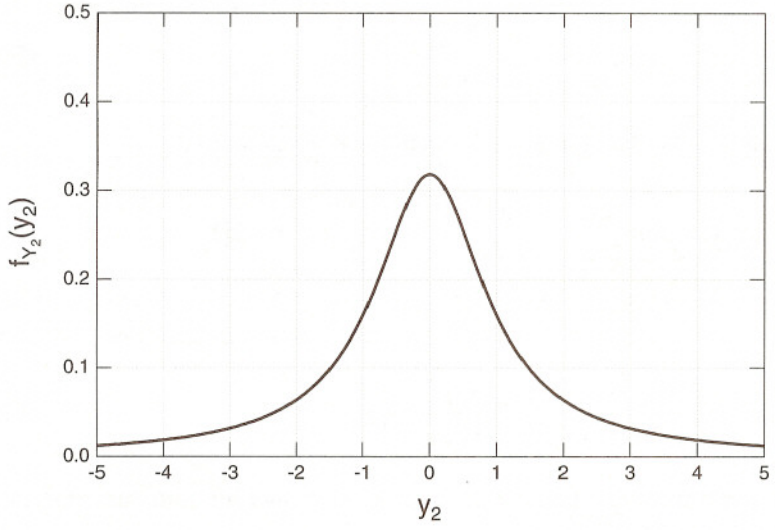


Figura 7.19 Densità di probabilità di Cauchy

7.4.7 Variabili aleatorie congiuntamente Gaussiane (vettori Gaussiani)

Consideriamo un vettore aleatorio $\mathbf{X} = [X_1, X_2, \ldots, X_n]^T$ costituito da n variabili aleatorie indipendenti. La densità di probabilità congiunta $f_{\mathbf{X}}(\mathbf{x})$ del vettore $\mathbf{X}$ è espressa, per l'indipendenza delle variabili, dal prodotto delle densità di tutte le componenti del vettore stesso, ovvero

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} f_{X_i}(x_i) \tag{7.4.49}$$

Se, inoltre, le variabili aleatorie sono Gaussiane, cioè $X_i \in \mathcal{N}(\eta_{X_i}, \sigma_{X_i}^2)$, $i = 1, 2, \ldots, n$, la funzione $f_{\mathbf{X}}(\mathbf{x})$ diventa

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_{X_i}^2}} e^{-\frac{(x_i-\eta_{X_i})^2}{2\sigma_{X_i}^2}} = \frac{1}{\sqrt{(2\pi)^n \prod_{i=1}^{n} \sigma_{X_i}^2}} e^{-\frac{1}{2}\sum_{i=1}^{n}\frac{(x_i-\eta_{X_i})^2}{2\sigma_{X_i}^2}} \tag{7.4.50}$$

Possiamo riscrivere questa densità usando il vettore dei valori medi $\boldsymbol{\eta}_{\mathbf{X}}$ (7.4.39) e la matrice di covarianza $\mathbf{C}_{\mathbf{X}}$ (7.4.41). Quest'ultima, per l'indipendenza (e quindi l'incorrelazione due a due) delle n variabili aleatorie, è *diagonale*:

$$\mathbf{C}_{\mathbf{X}} = \begin{bmatrix} \sigma_{X_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{X_2}^2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \sigma_{X_n}^2 \end{bmatrix}, \quad \det \mathbf{C}_{\mathbf{X}} = \prod_{i=1}^{n} \sigma_{X_i}^2 \tag{7.4.51}$$

per cui è facile riscrivere la densità di probabilità congiunta $f_{\mathbf{X}}(\mathbf{x})$ nel modo seguente:

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det \mathbf{C}_{\mathbf{X}}}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\eta}_{\mathbf{X}})^T \mathbf{C}_{\mathbf{X}}^{-1}(\mathbf{x}-\boldsymbol{\eta}_{\mathbf{X}})\right\} \tag{7.4.52}$$

Usciamo adesso dal caso *particolare* delle variabili indipendenti, e definiamo *Gaussiano* un vettore aleatorio $\mathbf{X}$ la cui densità di probabilità congiunta è espressa dalla (7.4.52) *qualunque sia la forma della matrice di covarianza* (purché ovviamente invertibile). Ribadiamo che la matrice $\mathbf{C}_{\mathbf{X}}$ è diagonale solo se le variabili aleatorie sono incorrelate. Tale caso *particolare* è solamente servito a scopo propedeutico, cioè per introdurre e motivare l'espressione *generale* (7.4.52) della densità di probabilità del vettore Gaussiano. Le variabili aleatorie $X_1, X_2, \dots, X_n$ che costituiscono tale vettore si dicono *congiuntamente Gaussiane* (o Gaussiane multivariate).

Un vettore Gaussiano $\mathbf{X} = [X_1, X_2, \dots, X_n]^T$ gode delle seguenti proprietà:

- il suo comportamento statistico è univocamente determinato dal suo vettore dei valori medi $\boldsymbol{\eta}_{\mathbf{X}}$ e dalla sua matrice di covarianza $\mathbf{C}_{\mathbf{X}}$;
- se le n variabili aleatorie costituenti $\mathbf{X}$ sono *incorrelate* a due a due, allora la densità di probabilità congiunta $f_{\mathbf{X}}(\mathbf{x})$ (7.4.52) può essere espressa come prodotto delle n densità di probabilità marginali $\{f_{X_i}(x_i), i = 1,2,\dots,n\}$; in altre parole, per un vettore Gaussiano, l'*incorrelazione* tra le variabili aleatorie implica la loro *indipendenza*;
- un qualunque sotto-vettore k-dimensionale ($k < n$) di $\mathbf{X}$ è ancora un insieme di variabili aleatorie congiuntamente Gaussiane. Segue (basta prendere $k = 1$)

che tutte le variabili aleatorie che compongono il vettore aleatorio Gaussiano sono marginalmente Gaussiane;

- il vettore aleatorio $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_m]^T$ generato a partire del vettore $\mathbf{X}$ con la trasformazione lineare

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \tag{7.4.53}$$

ove $\mathbf{A}$ è una matrice reale $m \times n$ e $\mathbf{b}$ è un vettore reale a n componenti, è anch'esso un vettore aleatorio Gaussiano con vettore dei valori medi

$$\boldsymbol{\eta}_Y = \mathbf{A}\,\boldsymbol{\eta}_X + \mathbf{b}$$

e matrice di covarianza

$$\mathbf{C}_Y = \mathbf{A}\,\mathbf{C}_X\,\mathbf{A}^T \tag{7.4.54}$$

- la densità congiunta di un qualunque sottogruppo di k variabili estratte dal vettore ($k < n$) *condizionata* a un qualunque (sotto)gruppo di m tra le restanti variabili ($m \le n - k$) è congiuntamente Gaussiana con opportuno vettore dei valori medi condizionati e matrice di covarianza condizionata.

Esempio 7.10

Il vettore aleatorio Gaussiano $\mathbf{V} = [X, Y, Z]^T$ ha vettore valori medi $\boldsymbol{\eta}_V = [0, 1, -1]^T$ e matrice di covarianza

$$\mathbf{C}_V = 4\begin{bmatrix} 1 & 0 & 0.5 \\ 0 & 1 & -0.5 \\ 0.5 & -0.5 & 1 \end{bmatrix} \tag{E7.10.1}$$

Determiniamo la probabilità che *la proiezione di* $\mathbf{V}$ *lungo la direzione* $\mathbf{d} = \left[0, 1/\sqrt{2}, 1/\sqrt{2}\right]^T$ *sia maggiore di* 2. Tale probabilità è data da

$$\Pr\{\mathbf{V} \cdot \mathbf{d} > 2\} = \Pr\left\{(Y + Z)/\sqrt{2} > 2\right\} = \Pr\left\{Y + Z > 2\sqrt{2}\right\} \tag{E7.10.2}$$

Essa può essere calcolata in maniera molto semplice se definiamo una nuova variabile aleatoria W come somma di Y e Z, ovvero

$$W = Y + Z \tag{E7.10.3}$$

La probabilità cercata è allora

$$\Pr\{\mathbf{V}\cdot\mathbf{d} > \sqrt{2}\} = \Pr\{W > 2\sqrt{2}\} \tag{E7.10.4}$$

La variabile aleatoria W, essendo combinazione lineare di due variabili aleatorie *congiuntamente Gaussiane*, è essa stessa Gaussiana. Il suo valor medio e la sua varianza, inoltre, sono calcolabili immediatamente attraverso le relazioni (E7.8.3)-(E7.8.4):

$$\eta_W = \eta_Y + \eta_Z = 1 - 1 = 0 \tag{E7.10.5}$$

e

$$\sigma_W^2 = \sigma_Y^2 + \sigma_Z^2 + c_{YZ} + c_{ZY} = 4(1 + 1 - 0.5 - 0.5) = 4 \tag{E7.10.6}$$

In conclusione, poiché abbiamo dimostrato che $W \in \mathcal{N}(0,4)$, la probabilità (E7.10.4) è pari a

$$\Pr\{\mathbf{V}\cdot\mathbf{d} > \sqrt{2}\} = \Pr\{W > 2\sqrt{2}\} = 1 - \Phi\left(\frac{2\sqrt{2}}{2}\right) \cong 0.0787 \tag{E7.10.7}$$

❑

7.4.9 Il teorema-limite centrale

Riprendiamo da un altro punto di vista la situazione dell'Esempio 7.8, consideriamo cioè la variabile aleatoria

$$Z_n = \sum_{i=1}^{n} X_i \tag{7.4.55}$$

che è ottenuta come somma delle n variabili X_i, $i = 1,\ldots,n$. Supponiamo poi che le n variabili siano *indipendenti* e che abbiano tutte *uguale funzione densità* $f_{X_i}(x) \equiv f(x)$ con valore atteso $\eta_{X_i} \equiv \eta$ e varianza $\sigma_{X_i}^2 \equiv \sigma^2$. Dai risultati del medesimo Esempio 7.8, sappiamo che il valore atteso η_n e la varianza σ_n^2 di Z_n sono rispettivamente

$$\eta_n = n\cdot\eta \quad , \quad \sigma_n^2 = n\cdot\sigma^2 \tag{7.4.56}$$

Se consideriamo un numero n di variabili X_i man mano *crescente*, vediamo che η_n e σ_n^2 crescono progressivamente con n. Definiamo allora, qualunque sia n, la versione *normalizzata* della variabile Z_n, precisamente la variabile aleatoria

$$S_n = \frac{Z_n - \eta_n}{\sigma_n} = \frac{Z_n - n \cdot \eta}{\sqrt{n} \cdot \sigma} \tag{7.4.57}$$

che, indipendentemente da n, ha comunque valore atteso *nullo* e varianza *unitaria*. Abbiamo già incontrato questo procedimento di normalizzazione nella definizione della variabile Gaussiana *standard* (7.3.31).

Un risultato fondamentale ("centrale") della teoria della probabilità è il **teorema-limite centrale** (Lyapunov). Sotto ipotesi relativamente poco restrittive sulla densità di probabilità $f_{X_i}(x) \equiv f(x)$ delle variabili X_i, si dimostra abbastanza facilmente che, al tendere di n all'infinito, la densità di probabilità $f_{S_n}(s)$ della variabile somma normalizzata S_n (7.4.57) *tende a una densità normale standard:*

$$\lim_{n \to \infty} f_{S_n}(s) = f_N(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} \tag{7.4.58}$$

In pratica, questo risultato dice che la somma di un gran numero di variabili aleatorie indipendenti segue con buona approssimazione una legge Gaussiana, a prescindere dalla particolare distribuzione di ciascuna di esse. Useremo il teorema-limite centrale nel prossimo capitolo quando modelleremo secondo una statistica Gaussiana un fenomeno fisico (il rumore termico) composto dalla sovrapposizione (somma) di moltissimi contributi elementari indipendenti. L'ipotesi di *equidistribuzione* delle variabili X_i permette di condurre una dimostrazione del teorema particolarmente semplice. Il teorema-limite centrale è però valido anche sotto ipotesi diverse e meno restrittive, sebbene la dimostrazione diventi più complessa.

Un'idea della tendenza verso una distribuzione Gaussiana della somma normalizzata (7.4.57) si può avere esaminando la Figura 7.20. Questa mostra le densità di probabilità delle variabili S_1 (a), S_2 (b), S_3 (c), S_{10} (d), S_{15} (e), $S_\infty = N$ (f), ottenute per ripetute convoluzioni e successiva normalizzazione a partire da una variabile *uniforme*:

$$f_{Z_n}(z) = \overbrace{f_X(z) \otimes f_X(z) \otimes \cdots \otimes f_X(z)}^{n-1 \text{ volte}} \quad , \quad f_{S_n}(s) = \sigma\sqrt{n} \cdot f_{Z_n}\left(\sigma\sqrt{n} \cdot s + n \cdot \eta\right) \tag{7.4.59}$$

La tendenza della densità n-esima verso una distribuzione Gaussiana è chiara; già per $n = 15$ la $f_{S_n}(s)$ e la $f_N(s)$ sono praticamente indistinguibili.

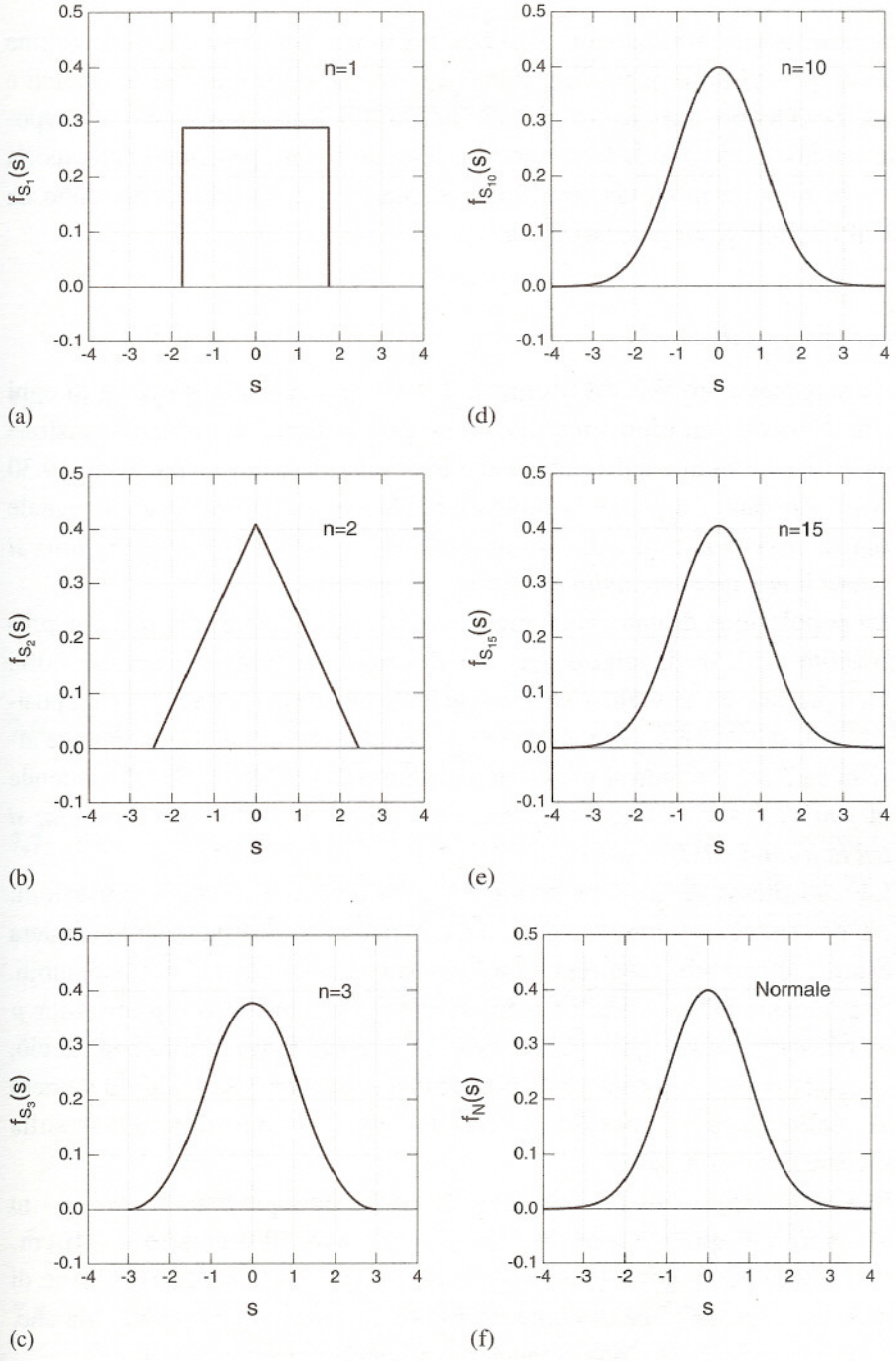


Figura 7.20 Dimostrazione grafica del teorema-limite centrale

Sommario

Contrariamente agli altri capitoli, non riassumeremo qui i risultati delle pagine precedenti, perché *tutto* questo capitolo è esso stesso un sommario di risultati e concetti che il lettore deve avere già fatto propri. Lo scopo di questa breve esposizione è stato quello di rivedere in modo sintetico i punti principali della teoria della probabilità, in modo da fissare uno scenario e le notazioni che fanno da sfondo al Capitolo 8 sui processi aleatori.

Esercizi proposti

7.1 La stazione radio WXYZ trasmette il segnale orario allo scoccare di ogni ora. L'ascoltatore-tipo sintonizza il proprio radioricevitore sulla stazione WXYZ a un istante uniformemente distribuito tra le ore 7.10 e le ore 19.30 nella giornata. Calcolare la probabilità che l'ascoltatore riceva il segnale orario entro 5 minuti dalla sintonizzazione su WXYZ (*Suggerimento: si adotti il minuto come unità di misura del tempo*).

7.2 La popolazione di una data regione è affetta dal virus Ebola con una probabilità dell'1%. Il miglior test per il virus ha affidabilità pari all'80%. Una persona viene scelta casualmente dalla popolazione data e risulta positiva al test. Qual è la probabilità che la persona scelta sia effettivamente affetta da Ebola? Come si modifica il risultato se l'affidabilità del test tende al 100%? E al 50%? Giustificare questi ultimi risultati (*Suggerimento: si usi la formula di Bayes*).

7.3 La studentessa XYZ viene sottoposta a un quiz con m risposte possibili. Se ha studiato l'argomento del quiz, ella risponderà certamente in maniera esatta, altrimenti sceglierà una risposta a caso tra le m disponibili. Supponiamo allora che XYZ abbia studiato l'argomento con probabilità p e che, sottoposta al quiz, abbia scelto la risposta esatta. Sulla base di ciò, qual è la probabilità che XYZ abbia studiato *davvero*? Se inoltre il numero m delle risposte possibili è grande, cosa si può concludere sulla preparazione di XYZ?

7.4 La scatola rappresentata in Figura 7.21 ha il fondo quadrato di lato $l = 1$ m al centro del quale è praticato un foro circolare di diametro $d = 10$ cm. Nella scatola vengono gettate a caso e indipendentemente 10 palline di piccolo diametro (cioè di diametro $<< d$). Determinare la probabilità che, alla fine della successione di lanci, nella scatola si trovino 7 palline.

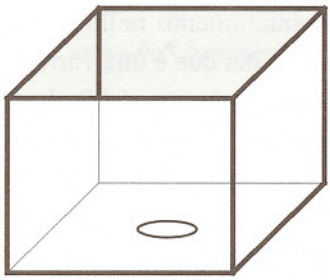


Figura 7.21

7.5 È data la variabile aleatoria X con densità di probabilità (esponenziale)

$$f_X(x) = \exp(-x)\, u(x)$$

Determinare e rappresentare la densità di probabilità della variabile aleatoria

$$Y = \sum_{i=0}^{+\infty} (-1)^i (X - i)\, \mathrm{rect}\big(x - (i + 1/2)\big)$$

7.6 È data la variabile aleatoria $X \in \mathcal{N}(0,1)$. Determinare la densità di probabilità $f_Y(y)$ e il valor medio η_Y della variabile aleatoria *lognormale* $Y = e^{-X}$.

7.7 Per ottenere con precisione il valore della potenza P dissipata su un resistore di resistenza nota r, si effettuano N operazioni di misura in condizioni indipendenti della tensione ai capi di detto resistore e si modellano i risultati di tali operazioni come N variabili aleatorie V_i, $i = 1,\ldots,N$, mutuamente indipendenti e uniformemente distribuite tra $-v$ e $+v$. Si forma quindi una stima della potenza dissipata come segue:

$$\hat{P} = \frac{1}{N} \sum_{i=1}^{N} \frac{V_i^2}{r}$$

Determinare il *valore atteso* $\eta_{\hat{P}}$ e l' *accuratezza* della stima, definita quest'ultima come la deviazione standard $\sigma_{\hat{P}}$ della variabile aleatoria $\hat{P}$. Commentare l'influenza sui risultati del numero N di misure effettuate.

7.8 In un esperimento aleatorio si fissano in modo indipendente due punti P_1 e P_2 su una retta. Le coordinate X_1 e X_2 rispettivamente di P_1 e P_2 sono modellate come variabili aleatorie indipendenti entrambe $\in \mathcal{N}(0,1)$.

Determinare il valor medio della distanza fra P_1 e P_2.

7.9 Due amici si danno appuntamento nella piazza del paese alle ore 0:00. L'ora di arrivo di ciascuno dei due è una variabile aleatoria uniformemente distribuita fra le ore 0:00 e le ore 1:00. Inoltre, i due amici arrivano indipendentemente l'uno dall'altro. Il primo arrivato aspetta l'altro per non più di 20 minuti e poi torna a casa. Calcolare la probabilità che i due amici si incontrino.

7.10 Su di un segmento di lunghezza a vengono scelti due punti casualmente e indipendentemente. Tali punti dividono il segmento dato in tre ulteriori segmenti di lunghezza $\leq a$. Calcolare la probabilità che il segmento "centrale" sia più lungo del segmento "a sinistra".

7.11 Sono assegnate due variabili aleatorie indipendenti X e Y, entrambe uniformemente distribuite fra –1 e 1. Trovare la probabilità che l'equazione di secondo grado in α

$$\alpha^2 + 2X\alpha + Y = 0$$

abbia radici reali.

7.12 Viene fissato a caso un punto su ciascuno dei due lati adiacenti di un quadrato. Calcolare la probabilità che l'area del triangolo formato dai lati del quadrato e dal segmento di retta congiungente i due punti sia minore di 1/8 dell'area del quadrato stesso.

7.13 Un libro A di 200 pagine e uno B di 300 vengono aperti indipendentemente e in modo casuale da due lettori. Determinare la probabilità che il numero della pagina alla quale il libro A viene aperto sia maggiore della pagina di apertura di B.

7.14 Per questioni di scarsa accuratezza nella manifattura, i valori della resistenza del resistore e della capacità del condensatore nel circuito di Figura 7.22 devono considerarsi come due variabili aleatorie rispettivamente R e C indipendenti e con densità di probabilità uniforme tra 0 e 2 kΩ per il resistore e tra 0 e 2 μF per il condensatore. Come indicato in figura, al condensatore inizialmente scarico viene applicata all'istante $t = 0$ una tensione pari a 1 V attraverso il resistore R. Calcolare la probabilità che la tensione ai capi del condensatore C all'istante $t_0 = 1$ ms sia inferiore a $1 - 1/e$ V.

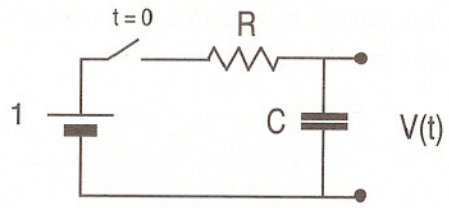


Figura 7.22

7.15 È data la variabile aleatoria X uniformemente distribuita tra -1 e 1, e la variabile Y indipendente da X e uniformemente distribuita tra 0 e 1. Determinare e rappresentare la funzione densità di probabilità $f_Z(z)$ della variabile aleatoria

$$Z = Y - |X|$$

7.16 Sono assegnate le due variabili aleatorie indipendenti X e Y, entrambe uniformemente distribuite fra 0 ed 1. Trovare la funzione densità di probabilità $f_Z(z)$ della variabile aleatoria

$$Z = \max(XY, X/Y)$$

7.17 È assegnato un gruppo di variabili aleatorie X_1, X_2, X_3, ... indipendenti e identicamente distribuite, con

$$f_{X_i}(x) = \frac{1}{2}\delta(x) + \frac{1}{2}\delta(x-1) \quad \forall i$$

Dato un generico numero intero n, trovare la probabilità p_n che valga congiuntamente

$$\begin{cases} X_1 + X_2 + \ldots + X_{n-1} \leq 5 \\ X_1 + X_2 + \ldots + X_{n-1} + X_n > 5 \end{cases}$$

7.18 Un nuotatore deve attraversare il fiume rappresentato schematicamente in Figura 7.23, e quindi si tuffa nel punto A e nuota con velocità di modulo aleatorio Y e direzione ortogonale alle sponde del fiume. Tuttavia, a causa dello scorrimento dell'acqua del fiume con velocità di modulo aleatorio X (e direzione indicata dalla freccia in figura), il nuotatore non approda nel punto B desiderato, bensì nel punto C. Sapendo che X e Y sono variabili aleatorie indipendenti e identicamente distribuite con densità esponenziale

unilatera di valor medio $\eta = 1$ m/s, e che la distanza fra le due sponde del fiume è $d = 100$ m, ricavare la densità di probabilità della variabile aleatoria S che rappresenta la distanza fra i punti B e C.

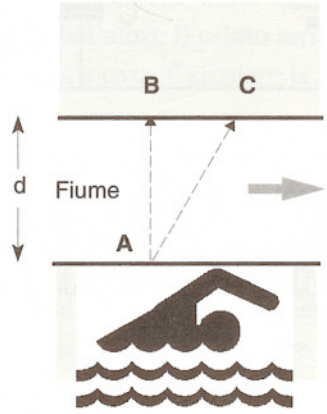


Figura 7.23

7.19 Sono assegnate due variabili aleatorie indipendenti X e Y, entrambe uniformemente distribuite fra 0 e 1. Si consideri quindi il sistema di variabili aleatorie

$$\begin{cases} Z = X - Y \\ V = X + Y \end{cases}$$

Trovare la densità di probabilità condizionata $f_{Z|A}(Z|A)$, ove $A \equiv \{V \leq 1\}$.

7.20 Dimostrare che la densità di probabilità congiunta di due variabili aleatorie X e Y congiuntamente Gaussiane si può esprimere come segue:

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} e^{-\frac{\left(\frac{x-\eta_X}{\sigma_X}\right)^2 - 2\rho_{XY}\left(\frac{x-\eta_X}{\sigma_X}\right)\left(\frac{y-\eta_Y}{\sigma_Y}\right) + \left(\frac{y-\eta_Y}{\sigma_Y}\right)^2}{2\left(1-\rho_{XY}^2\right)}}$$

ove i 5 parametri η_X, η_Y, σ_X, σ_Y, ρ_{XY} sono rispettivamente i due valori attesi, le due varianze e il coefficiente di correlazione delle due variabili date.

8

Segnali aleatori

8.1 Dai segnali determinati ai segnali aleatori

Già nel Capitolo 1 abbiamo introdotto due importanti classi distinte di segnali: *determinati* o *aleatori*. L'esempio tipico di segnale determinato è l'"onda quadra" prodotta da un generatore di forme d'onda elettronico (Figura 8.1): di questo segnale è possibile conoscere *a priori* l'andamento, perché è possibile controllarne l'ampiezza picco-picco $2A$ e il periodo di ripetizione T_0 (o la frequenza fondamentale f_0) agendo sui controlli dello strumento.

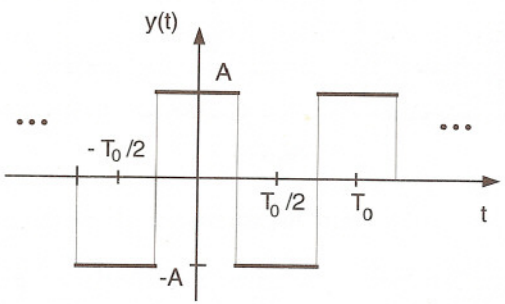


Figura 8.1 Esempio di segnale determinato: l'*onda quadra*

Consideriamo invece la tensione raccolta tra due elettrodi posti sul corpo di un paziente per la misura di un *elettrocardiogramma*. A seconda del paziente e del particolare stato di salute in cui egli si trova, si ottengono diverse forme d'onda, come quelle mostrate in Figura 8.2.

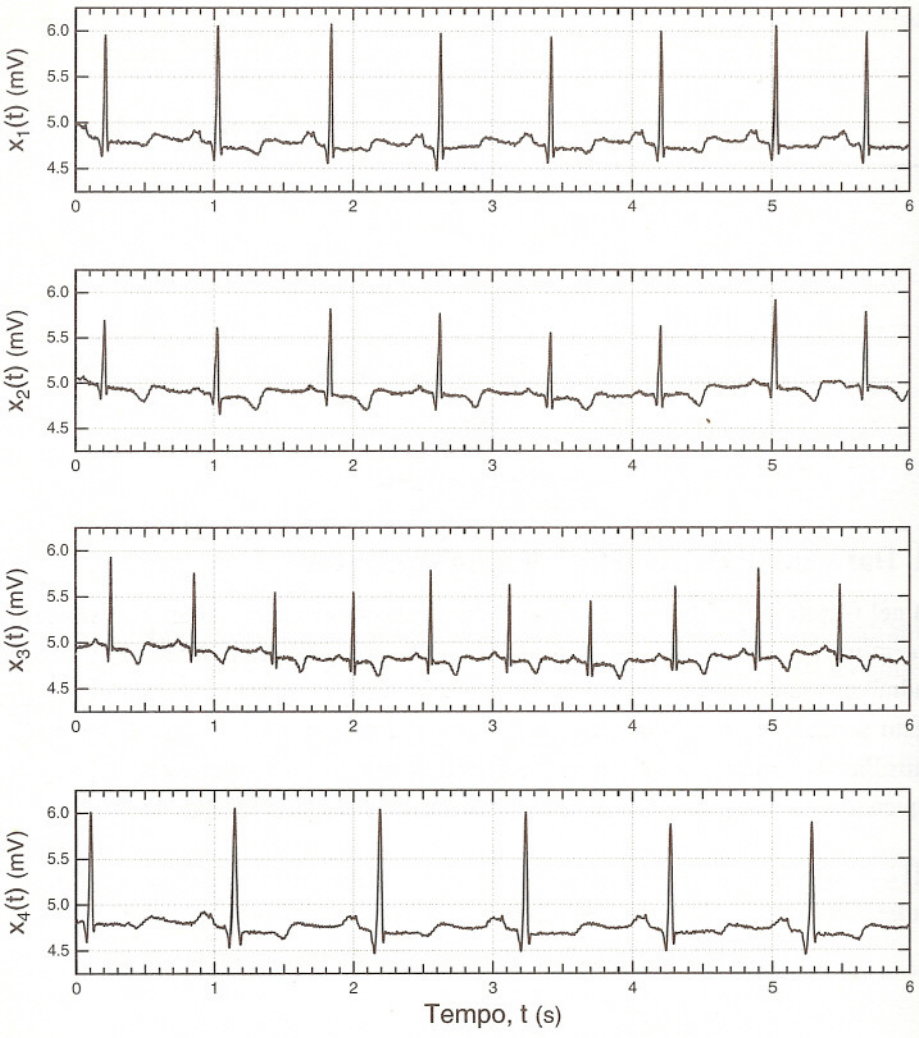


Figura 8.2 Esempi di elettrocardiogramma in pazienti affetti da aritmia

Queste forme d'onda non sono ovviamente predicibili a priori (come l'onda quadra), né sul breve né sul lungo termine. Sull'andamento di questi segnali si hanno solo informazioni generiche: come si nota dalla Figura 8.2, l'ampiezza dell'elettrocardiogramma è di pochi mV attorno a un valore continuo diverso da zero (informazione sulle *ampiezze*), e il segnale mostra una certa periodicità con un periodo valutabile all'incirca tra 0.5 e 1 s (informazione sull'*evoluzione temporale*). L'unica maniera di conoscere l'andamento di un elettrocardiogramma è

quella di *osservarlo* e, come quotidianamente in uso nella pratica medica, *registrarlo* sotto forma di grafico su di una striscia di carta, o, negli ambulatori più moderni, sotto forma di un *file* in uno strumento elettronico digitale. Il segnale non è dunque determinato, cioè non è predicibile, ed è noto solo *a posteriori*: in una parola, abbiamo a che fare con un segnale *aleatorio*. Nonostante questa impossibilità di effettuare una predizione, dovremo comunque essere in grado di studiare i segnali aleatori, per poter progettare sistemi che osservano ed elaborano correttamente tali segnali (ad esempio un buon elettrocardiografo).

Lo strumento matematico per eccellenza che permette di studiare i segnali aleatori è naturalmente la *teoria della probabilità*, i cui concetti basilari sono stati rivisti brevemente nel Capitolo 7. La modellizzazione di un segnale aleatorio viene effettuata attraverso la *teoria dei processi aleatori* (o *stocastici*) che in pratica costituisce l'oggetto di tutto questo capitolo.

8.1.1 Definizione di processo aleatorio

Torniamo dunque all'esempio dell'elettrocardiogramma, e supponiamo di compiere uno dei soliti "esperimenti ideali". Ammettiamo di essere in grado di raccogliere tutti gli esseri umani del mondo in un ambulatorio e di misurare e registrare i loro rispettivi elettrocardiogrammi in un "enorme archivio". Dovremmo cioè aggiungere ai quattro segnali $x_1(t), x_2(t), x_3(t), x_4(t)$ di Figura 8.2, chiamati *funzioni campione*, i restanti 5 miliardi circa di funzioni campione di tutti gli altri esseri umani. Ogni volta che in un ambulatorio di cardiologia si presentasse un paziente, l'osservazione del suo elettrocardiogramma si ridurrebbe quindi alla *selezione del segnale appropriato nell'archivio*. Naturalmente, non sapendo a priori *chi* si presenterà in ambulatorio, non è possibile sapere che forma avrà il segnale finche il paziente non si sarà presentato (cioè finché non sarà stato effettivamente *misurato* l'elettrocardiogramma): in questo senso il segnale è *aleatorio*.

Questo che abbiamo chiamato "esperimento ideale" è una definizione un po' romanzata di *processo aleatorio*. Questa definizione richiede innanzitutto di considerare un *esperimento aleatorio* o, meglio, uno *spazio di probabilità* caratterizzato da uno spazio campione $\Omega = \{\omega_i\}$ (qui per semplicità discreto), da una classe degli eventi S e dalla legge di probabilità $\Pr(\cdot)$ definita su S. Si deve poi individuare un insieme di *funzioni del tempo* $x_i(t)$ (le *funzioni campione*) in numero pari a quello dei risultati dell'esperimento ω_i. Infine, si deve *istituire una corrispondenza* che associa a ciascun risultato ω_i dell'esperimento una delle possibili funzioni campione $x_i(t)$:

$$X(\omega_i;t) = x_i(t) \tag{8.1.1}$$

Questa corrispondenza, rappresentata in Figura 8.3, costituisce appunto il *processo aleatorio*. Quando si effettua una *prova* dell'esperimento (una misura di un elettrocardiogramma), si ha un *risultato* dello spazio campione (un paziente) cui è associata una funzione campione (un elettrocardiogramma), cioè il segnale che effettivamente viene osservato, e che prende il nome di *realizzazione* del processo[1].

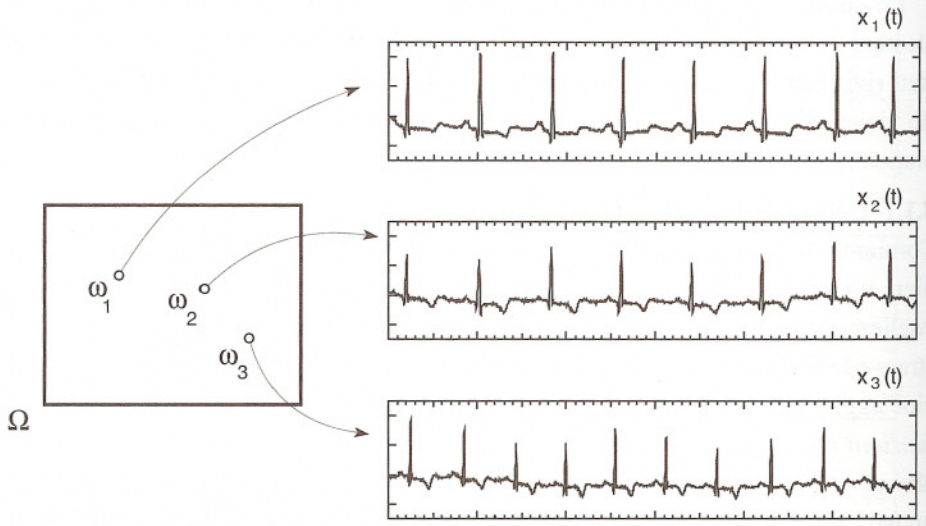


Figura 8.3 Rappresentazione grafica della definizione di processo aleatorio

Il processo aleatorio caratterizzato dalla (8.1.1) si indica comunemente con $X(t)$ omettendo per semplicità la dipendenza dal risultato ω_i dello spazio campione Ω. Tale dipendenza, come nel caso delle variabili aleatorie, deve essere sempre considerata implicita (si veda il Paragrafo 7.3.1).

Discutiamo adesso le conseguenze di questa definizione. Come è già stato chiarito, fissare nel processo aleatorio $X(\omega_i;t)$ il risultato dell'esperimento, ad esempio ω_1, significa selezionare quella tra le varie funzioni campione che si è *realizzata* in una data prova; non c'è più alcuna aleatorietà e il processo diventa *a posteriori* il segnale *determinato* $X(\omega_1;t)$, cioè la funzione campione $x_1(t)$.

[1] Normalmente, *realizzazione* viene usata come sinonimo di *funzione campione*. In realtà, le funzioni campione sono tutti i *possibili* segnali del processo, mentre la *realizzazione* è quello tra i possibili segnali che viene effettivamente osservato in una data prova dell'esperimento.

Viceversa, cosa succede se fissiamo arbitrariamente un certo *istante di tempo* t_1 nel processo $X(\omega_i;t)$? Questa operazione è rappresentata nella Figura 8.4, in cui sono riportate sullo stesso grafico quattro funzioni campione di un processo dato e in cui viene evidenziato l'istante temporale $t = t_1$. Come si nota dal grafico, il valore del processo $X(\omega_i;t_1)$ per un istante fissato è un insieme di (quattro) valori ottenuti "campionando" le (quattro) funzioni campione a quell'istante. Ogni valore risulta automaticamente corrispondente a un risultato dello spazio campione (ovviamente, quello della relativa funzione campione): in una parola, il valore del processo a un dato istante è una *variabile aleatoria*. Questa conclusione si accorda bene con il concetto elementare che si può avere di un segnale aleatorio, cioè di un segnale il cui valore a un dato istante non sia esattamente determinabile.

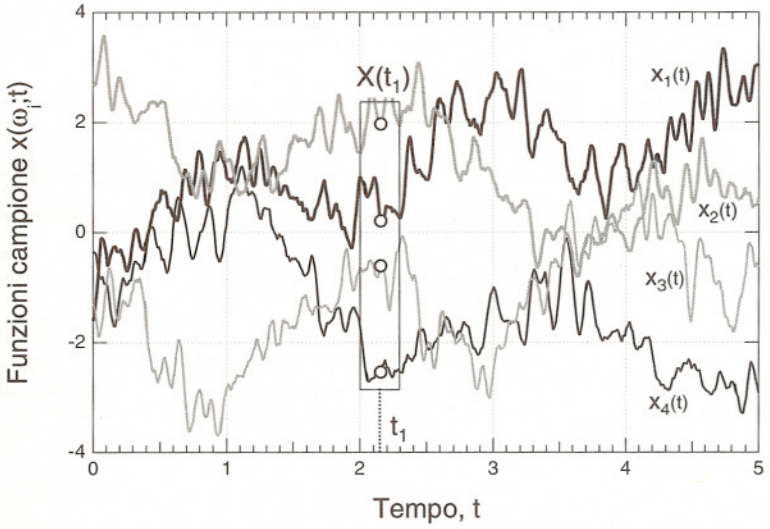


Figura 8.4 Variabile aleatoria estratta da un processo

8.1.2 Processi parametrici

Un semplice esempio di processo aleatorio è il seguente:

$$X(\omega;t) = e^{-A(\omega)t}\,\mathrm{u}(t) \tag{8.1.2}$$

in cui $A(\omega)$ è una variabile aleatoria con distribuzione uniforme nell'intervallo $(0,1/T)$. Omettendo la dipendenza dal risultato ω si ha equivalentemente

$$X(t) = e^{-At}\,\mathrm{u}(t) \tag{8.1.3}$$

Questo è un esempio di processo *parametrico*: viene definita una classe di funzioni campione il cui andamento dipende dal valore di un numero finito di *variabili aleatorie* (parametri). Queste variabili aleatorie servono in un certo senso da "intermediario" tra lo spazio campione e le funzioni campione: l'associazione tra i risultati dell'evento e la funzione campione non è *diretta* come in Figura 8.3, bensì passa attraverso il particolare valore che la variabile aleatoria prende nella prova dell'esperimento, come in Figura 8.5. Alcune funzioni campione $x(\omega;t)$ del processo dato corrispondenti a valori diversi della variabile aleatoria A sono rappresentate nella Figura 8.6.

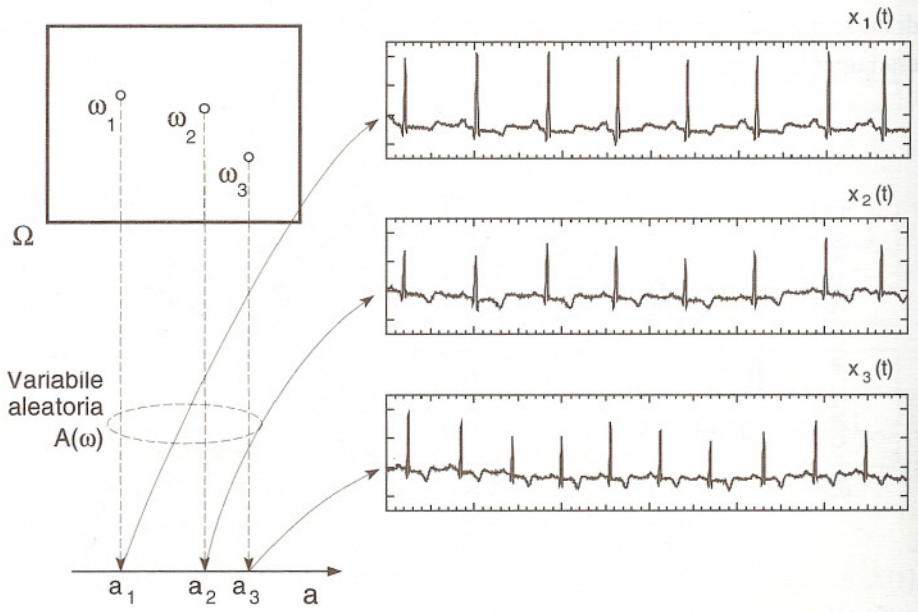


Figura 8.5 Costruzione di un processo parametrico

Come ulteriore esempio di processo parametrico, consideriamo l'oscillazione sinusoidale prodotta da un generatore di forme d'onda elettronico. Di questa oscillazione possono essere controllate l'ampiezza e la frequenza, mentre lo strumento non consente in genere di controllarne la *fase iniziale*. Il modello più appropriato di questo segnale è allora il processo parametrico

$$X(t) = a\cos(2\pi f_0 t + \Theta) \tag{8.1.4}$$

in cui a ed f_0 sono quantità note, mentre la fase iniziale Θ è una variabile aleatoria uniformemente distribuita in $[0,2\pi)$ (cioè è completamente casuale).

Torneremo su quest'ultimo processo nell'Esempio 8.2.

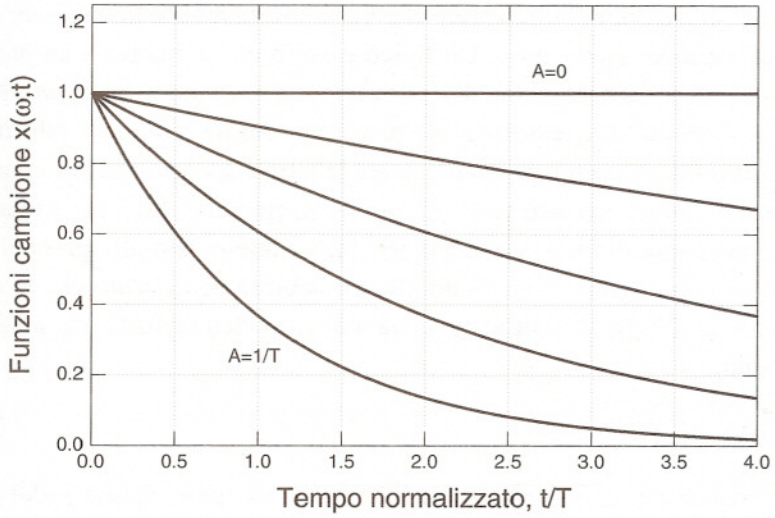


Figura 8.6 Funzioni campione del processo parametrico (8.1.4)

8.1.3 Caratterizzazione statistica di un processo aleatorio

È chiaro a questo punto della discussione che, contrariamente al caso di un segnale determinato, non ha senso parlare dell'*andamento* di un processo. D'altronde, l'elencazione esaustiva di tutte le funzioni campione del processo e, soprattutto, la loro associazione ai risultati dello spazio campione è un procedimento impensabile nella grande maggioranza dei casi (eccezione notevole, i processi parametrici). Si pone quindi il problema della *caratterizzazione* delle proprietà di un processo dal punto di vista *statistico*.

Cominciamo con l'osservare che, stante la discussione al paragrafo precedente, se si fissa un arbitrario istante di tempo $t = t_1$, il valore del processo $X(t_1)$ a quell'istante è in generale una *variabile aleatoria*. Come è stato richiamato nel Paragrafo 7.3, il comportamento statistico di questa variabile può essere descritto mediante la sua funzione distribuzione. È dunque ragionevole, in virtù di queste considerazioni, definire la *funzione distribuzione di probabilità del primo ordine del processo* mediante la relazione

$$F_X(x;t_1) \triangleq \Pr\{X(t_1) \le x\} \tag{8.1.5}$$

Naturalmente questa funzione dipende anche da una variabile *temporale* perché le proprietà statistiche della variabile aleatoria $X(t_1)$ cambiano, in generale, al

cambiare dell'istante di tempo t_1 al quale si "campiona" il processo.

Potrebbe sembrare che la funzione $F_X(x;t_1)$ sia sufficiente a caratterizzare le proprietà del processo. Per dimostrare che quest'affermazione *non è vera*, consideriamo un esempio elementare. Un tipico caso in cui si ricorre a un processo aleatorio per modellare una serie di osservazioni è la *quotazione di un titolo in borsa*. L'andamento di questo segnale al variare del tempo non è (sfortunatamente) prevedibile, e quindi possiamo usare la teoria dei processi per cercare di ottimizzare i nostri investimenti in borsa. Chiamata $X(t)$ la quotazione (aleatoria) del titolo di interesse, un investitore è interessato alla probabilità di realizzare un utile, ossia è interessato alla probabilità dell'*evento* che la quotazione del titolo all'istante t_2 di vendita sia maggiore della quotazione all'istante t_1 di acquisto:

$$\Pr\{X(t_2) \geq X(t_1)\} \tag{8.1.6}$$

Questa probabilità *non* può essere ricavata dalla funzione $F_X(x;t_1)$ (del primo ordine) perché richiede la considerazione *congiunta* di *due* variabili aleatorie estratte dallo stesso processo *in istanti distinti*. È quindi necessario introdurre una descrizione del *secondo ordine* del processo. Si devono considerare cioè due istanti di osservazione t_1 e t_2 estraendo, così, due variabili aleatorie $X(t_1)$ e $X(t_2)$. Il comportamento statistico di questa *coppia* di variabili aleatorie è allora completamente descritto dalla loro funzione di distribuzione *congiunta*, cioè dalla funzione

$$F_X(x_1,x_2;t_1,t_2) \triangleq \Pr\{X(t_1) \leq x_1, X(t_2) \leq x_2\} \tag{8.1.7}$$

nota come *distribuzione di probabilità del secondo ordine del processo.*

Come è chiaro, il ragionamento che ha portato a estendere la descrizione dal primo al secondo ordine può essere iterato a piacere. La conclusione è che la descrizione statistica di un processo è *completa* solo quando si è in grado di caratterizzare il comportamento statistico *congiunto* di un numero n arbitrario di variabili aleatorie $X(t_1), X(t_2), \ldots, X(t_n)$ estratte da $X(t)$ a n istanti diversi, comunque grande sia il numero intero n, e comunque si scelga la n-upla di istanti $(t_1,t_2,\ldots,t_n)$. Ciò richiede, quindi, la conoscenza della *funzione distribuzione di probabilità dell'n-esimo ordine* definita dalla relazione:

$$F_X(x_1,x_2,\ldots,x_n;t_1,t_2,\ldots,t_n) \triangleq \Pr\{X(t_1) \leq x_1, X(t_2) \leq x_2, \ldots, X(t_n) \leq x_n\} \tag{8.1.8}$$

per ogni valore del parametro n e per ogni valore del vettore $(t_1,t_2...,t_n)$. Una descrizione completa del tutto equivalente (e anche più usata in pratica) si basa sulla *funzione densità di probabilità di ordine n* del processo:

$$f_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n)\stackrel{\Delta}{=}\frac{\partial^n F_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n)}{\partial x_1\partial x_2...\partial x_n} \tag{8.1.9}$$

Riassumendo, la caratterizzazione statistica completa del processo aleatorio $X(t)$ richiede la conoscenza della *classe* di funzioni densità di probabilità $f_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n)$ per qualunque numero n di variabili aleatorie $X(t_1)$, $X(t_2)$,..., $X(t_n)$, cioè di qualunque *ordine*.

Osserviamo che ciò che è stato introdotto come una *caratterizzazione* può anche essere usato come *definizione* alternativa di processo aleatorio: *un processo aleatorio è una famiglia di variabili aleatorie dipendenti dalla variabile temporale t e caratterizzate dalla classe delle densità di probabilità congiunte* $f_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n)$. In questo testo abbiamo preferito introdurre il processo aleatorio come associazione di funzioni campione ai risultati dello spazio campione perché tale definizione è più orientata allo studio del processo come *segnale*.

È abbastanza chiaro che, assegnato un processo $X(t)$, è impresa disperata pervenire alla sua conoscenza statistica completa. In alcuni casi è sufficiente in realtà conoscere la distribuzione (o densità) di probabilità del primo ordine, raramente si cerca di misurare o calcolare quella del secondo. Molto più spesso ci si accontenta di *parametri statistici semplificati* (si veda a questo proposito il Paragrafo 7.3.4) che saranno discussi nel prossimo paragrafo.

Esempio 8.1

Un punto materiale si muove su una linea retta con moto uniforme. Di questo punto si misurano con incertezza la posizione e la velocità all'istante iniziale $t=0$. Queste due grandezze vengono dunque modellate rispettivamente con una variabile aleatoria $S_0\in\mathcal{U}(-1,1)$, e una seconda variabile $V_0\in\mathcal{N}(0,4/\mathrm{T}^2)$ indipendente da S_0. Indichiamo con $S(t)$ il *processo aleatorio* che rappresenta la posizione del punto al generico istante t e determiniamo la densità di probabilità del primo ordine $f_S(s;t_1)$ di tale processo.

Il segnale $S(t)$ è un *processo aleatorio parametrico*, essendo espresso dalla seguente relazione:

$$S(t) = S_0 + V_0 t \tag{E8.1.1}$$

Le funzioni campione di questo processo sono *rette* di pendenza e intercetta sugli assi aleatorie. Fissiamo dunque il *particolare* istante di tempo t_1, ed estraiamo dal processo la variabile aleatoria

$$S(t_1) = S_0 + V_0 t_1 \tag{E8.1.2}$$

Per determinare la densità $f_S(s;t_1)$ cominciamo con l'osservare che il prodotto $S_1 \equiv V_0 t_1$ è una variabile aleatoria $\in \mathcal{N}(0, 4t_1^2/\mathrm{T}^2)$ ancora indipendente da S_0. La variabile estratta dal processo è

$$S(t_1) = S_0 + S_1 \tag{E8.1.3}$$

e la funzione $f_S(s;t_1)$ è data, quindi, dalla *convoluzione* tra la densità di probabilità $f_{S_0}(s)$ di S_0 e quella $f_{S_1}(s)$ di S_1:

$$f_S(s;t_1) = f_{S_0}(s) \otimes f_{S_1}(s) = \int_{-\infty}^{+\infty} f_{S_0}(u) f_{S_1}(s-u)\, du \tag{E8.1.4}$$

ove naturalmente il parametro t_1 è "nascosto" in $f_{S_1}(s_1)$. Ricordando che:

$$f_{S_0}(s) = \frac{1}{2}\mathrm{rect}\left(\frac{s}{2}\right) \tag{E8.1.5a}$$

$$f_{S_1}(s) = \frac{T}{|t_1|\sqrt{8\pi}} e^{-\frac{s^2 T^2}{8t_1^2}} \tag{E8.1.5b}$$

dalla (E8.1.4) si ricava che

$$f_S(s;t_1) = \int_{-\infty}^{\infty} \frac{1}{2}\mathrm{rect}\left(\frac{u}{2}\right) \frac{T}{|t_1|\sqrt{8\pi}} e^{-\frac{(s-u)^2 T^2}{8t_1^2}}\, du = \frac{1}{2}\int_{-1}^{1} \frac{T}{|t_1|\sqrt{8\pi}} e^{-\frac{(s-u)^2 T^2}{8t_1^2}}\, du \tag{E8.1.6}$$

Effettuando il cambiamento di variabile $\alpha = (u-s)T/(2|t_1|)$ si ottiene:

$$f_S(s;t_1) = \frac{1}{2}\int_{-(1+s)T/(2|t_1|)}^{(1-s)T/(2|t_1|)} \frac{1}{\sqrt{2\pi}} e^{-\frac{\alpha^2}{2}}\, d\alpha = \frac{1}{2}\left[\Phi\left(\frac{(1-s)T}{2|t_1|}\right) - \Phi\left(-\frac{(1+s)T}{2|t_1|}\right)\right] \tag{E8.1.7}$$

ove $\Phi(x)$ è la funzione distribuzione di una variabile aleatoria normale standard (7.3.3). L'andamento della funzione $f_S(s;t_1)$ è illustrato nella Figura 8.7 per i valori indicati del parametro t_1. Il lettore spieghi perché questa funzione tende a

diventare "squadrata" man mano che t_1 si avvicina a zero.

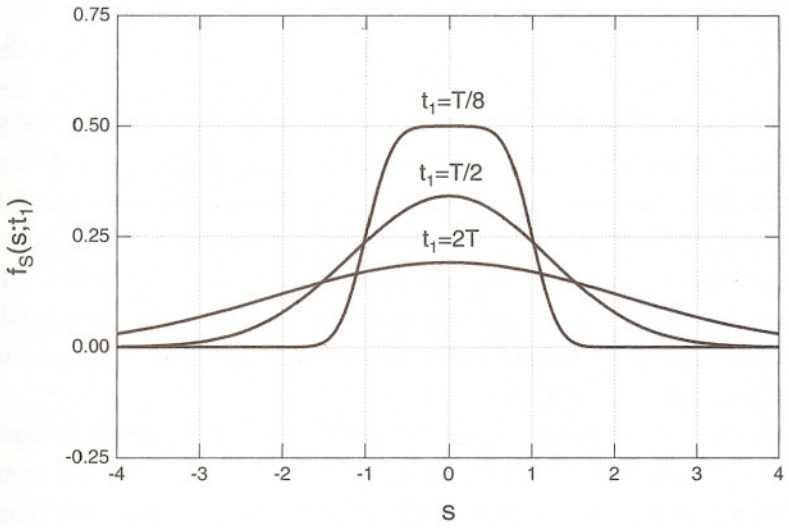


Figura 8.7 Densità di probabilità del primo ordine del processo dell'Esempio 8.1

❑

8.2 Indici statistici del 1° e 2° ordine di un processo aleatorio

8.2.1 Funzioni valor medio, potenza, varianza

Come già accennato nel paragrafo precedente, la caratterizzazione statistica completa del processo aleatorio $X(t)$ richiede la conoscenza della *classe* di funzioni densità di probabilità $f_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n)$, ed è quindi estremamente complessa da ottenere. Il comportamento del processo è spesso abbastanza ben descritto da indici statistici *semplificati*.

Da questo punto di vista, una grandezza particolarmente significativa nella descrizione statistica semplificata di un processo aleatorio $X(t)$ è la sua funzione *valor medio statistico* $\eta_X(t)$. Per definizione, il valore di questa funzione a un istante assegnato $t=\bar{t}$ è il valor medio della variabile aleatoria $X(\bar{t})$, estratta dal processo all'istante stesso:

$$\eta_X(\bar{t}) = E\{X(\bar{t})\} = \int_{-\infty}^{+\infty} x\, f_X(x;\bar{t})\, dx \qquad (8.2.1)$$

Al variare di $\bar{t}$ si generano infinite variabili aleatorie, ciascuna con un diverso valor medio. Ripetendo il calcolo nella (8.2.1) per ogni valore della variabile

temporale si ricava l'andamento della *funzione* $\eta_X(t)$ definita come segue:

$$\eta_X(t) \triangleq E\{X(t)\} = \int_{-\infty}^{+\infty} x\, f_X(x,t)\, dx \tag{8.2.2}$$

La funzione valor medio (8.2.2) rappresenta una *statistica del primo ordine* di $X(t)$ poiché il suo calcolo prevede la considerazione di *una sola* variabile aleatoria estratta dal processo, e quindi richiede la conoscenza della sola densità di probabilità del primo ordine del processo stesso. La dipendenza della funzione $\eta_X(t)$ dalla variabile tempo è dovuta al fatto che la densità di probabilità $f_X(x;t)$ del primo ordine di $X(t)$ è, usualmente, una funzione di tale variabile perché le proprietà statistiche della variabile aleatoria estratta dal processo cambiano al variare del tempo (si veda la Figura 8.7).

La funzione $\eta_X(t)$ rappresenta una sorta di "compendio" dell'andamento di tutte le funzioni campione del processo, pesate ciascuna con la relativa probabilità di presentazione, e per questo *non necessariamente* è una delle funzioni campione del processo $X(t)$ (si veda la Figura 8.8).

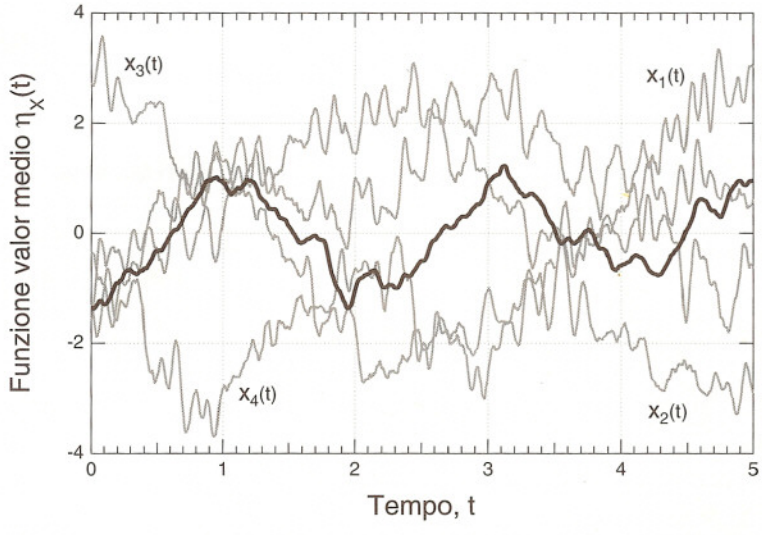


Figura 8.8 Funzione valor medio statistico $\eta_x(t)$ di un processo aleatorio $X(t)$

Esempio 8.2

Prendiamo in considerazione una versione *modificata* del processo aleatorio parametrico introdotto alla fine del Paragrafo 8.1.1:

$$X(t) = a\cos(2\pi f_0 t + \Theta) \tag{E8.2.1}$$

ove a ed f_0 sono noti, mentre la fase iniziale stavolta è $\Theta \in \mathcal{U}(0,\pi)$. Le funzioni campione del processo sono segnali cosinusoidali aventi la medesima frequenza f_0, ma una diversa fase iniziale Θ appartenente all'intervallo $[0,\pi]$. Se immaginiamo di *fissare* il valore del tempo t, abbiamo una variabile aleatoria $a\cos(2\pi f_0 t + \Theta)$ ottenuta come *trasformazione* della variabile Θ. Per ricavare il valor medio di tale variabile possiamo quindi usare il *teorema del valor medio* (7.3.24):

$$E\{X(t)\} = E_\Theta\{a\cos(2\pi f_0 t + \Theta)\} = \int_{-\infty}^{+\infty} a\cos(2\pi f_0 t + \theta) f_\Theta(\theta)\, d\theta$$

$$= \frac{a}{\pi}\int_0^\pi \cos(2\pi f_0 t + \theta)\, d\theta = \frac{a}{\pi}\left[\sin(2\pi f_0 t + \pi) - \sin(2\pi f_0 t)\right] = -\frac{2a}{\pi}\sin(2\pi f_0 t) \tag{E8.2.2}$$

ove il simbolo $E_\Theta\{\cdot\}$ rappresenta l'operazione di media statistica rispetto alla variabile Θ. La Figura 8.9 mostra alcune funzioni campione del processo insieme con la funzione valor medio $\eta_X(t)$ appena ricavata:

$$\eta_X(t) = -\frac{2a}{\pi}\sin(2\pi f_0 t) \tag{E8.2.3}$$

Il lettore ricavi la funzione valor medio del processo nella versione del Paragrafo 8.1.1, in cui cioè $\Theta \in \mathcal{U}(0,2\pi)$. ❑

Un'altra grandezza statistica *del primo ordine* utile per caratterizzare statisticamente il processo $X(t)$ è la funzione *potenza media statistica istantanea* $P_X(t)$ (brevemente, *potenza media*):

$$P_X(t) \triangleq E\{X^2(t)\} = \int_{-\infty}^{+\infty} x^2 f_X(x;t)\, dx \tag{8.2.3}$$

che è il diretto analogo, per i segnali aleatori, della potenza istantanea $p_x(t) = x^2(t)$ dei segnali determinati. Per i segnali aleatori si definisce anche la *funzione varianza* del processo:

$$\sigma_X^2(t) \triangleq E\{(X(t) - \eta_X(t))^2\} = \int_{-\infty}^{+\infty} (x - \eta_X(t))^2 f_X(x;t)\, dx \tag{8.2.4}$$

cioè la varianza della variabile aleatoria ottenuta fissando l'istante t. Dalla (8.2.4) si ricava facilmente l'uguaglianza

$$\sigma_X^2(t) = P_X(t) - \eta_X^2(t) \tag{8.2.5}$$

che esprime la varianza di $X(t)$ in funzione del suo valor medio e della sua potenza media statistica (si veda anche la (7.3.29)).

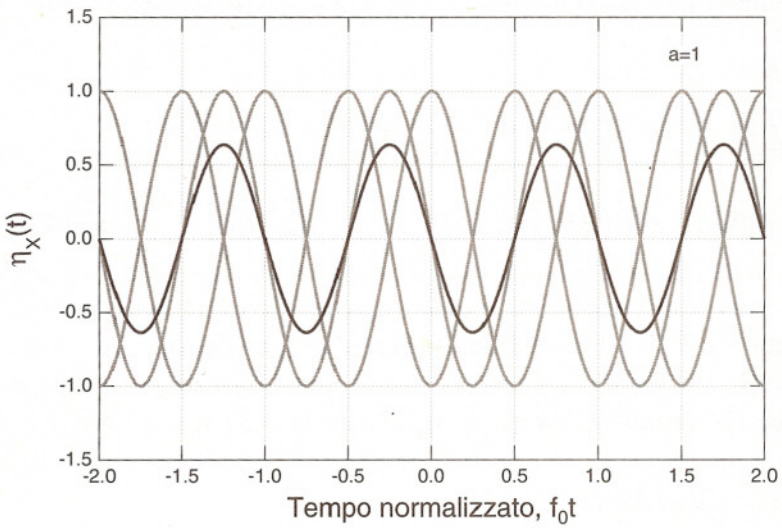


Figura 8.9 Funzioni campione e funzione valor medio del processo dell'Esempio 8.2

8.2.2 Funzioni di autocorrelazione e autocovarianza

Introduciamo adesso due parametri statistici del *secondo ordine* di fondamentale importanza per lo studio dei segnali aleatori, rimandando di qualche pagina la discussione sul *significato* e sull'utilità dei medesimi (che apparirà chiara nel contesto dei processi *stazionari*).

Fissiamo adesso *due* istanti di tempo arbitrari t_1 e t_2 sul nostro processo, ottenendo le due variabili aleatorie $Y = X(t_1)$ e $Z = X(t_2)$. È significativo allora calcolare la *correlazione* $r_{YZ} = E\{YZ\}$ (7.4.27) fra queste due variabili. Naturalmente, il valore di questa correlazione risulterà *funzione dei due istanti* t_1 e t_2 ai quali le variabili sono state estratte, e potrà essere calcolata solo conoscendo la funzione densità di probabilità *del secondo ordine* del processo:

$$R_X(t_1,t_2) \triangleq E\{X(t_1)X(t_2)\} = \int_{x_1=-\infty}^{+\infty} \int_{x_2=-\infty}^{+\infty} x_1 x_2 f_X(x_1,x_2;t_1,t_2)\,dx_1\,dx_2 \tag{8.2.6}$$

La $R_X(t_1,t_2)$ si chiama *funzione di autocorrelazione* perché le due variabili aleatorie di cui si calcola la correlazione sono estratte dallo *stesso* processo aleatorio.

Se invece tra le due variabili aleatorie $X(t_1)$ e $X(t_2)$ calcoliamo la *covarianza* (7.4.28) otteniamo la *funzione di autocovarianza* $C_X(t_1,t_2)$ di $X(t)$:

$$C_X(t_1,t_2) \triangleq E\{[X(t_1)-\eta_X(t_1)][X(t_2)-\eta_X(t_2)]\}$$
$$= \int_{x_1=-\infty}^{+\infty}\int_{x_2=-\infty}^{+\infty}[x_1-\eta_X(t_1)]\cdot[x_2-\eta_X(t_2)]f_X(x_1,x_2;t_1,t_2)dx_1\,dx_2 \tag{8.2.7}$$

Da questa definizione si ricava immediatamente la relazione:

$$C_X(t_1,t_2) = R_X(t_1,t_2)-\eta_X(t_1)\eta_X(t_2) \tag{8.2.8}$$

che esprime la funzione di autocovarianza di $X(t)$ attraverso le funzioni valor medio e autocorrelazione.

Esempio 8.3

Ricaviamo la funzione di autocorrelazione $R_X(t_1,t_2)$ del processo dell'Esempio 8.2:

$$X(t) = a\cos(2\pi f_0 t+\Theta) \tag{E8.3.1}$$

Osservando il processo aleatorio agli istanti t_1 e t_2 otteniamo rispettivamente le variabili aleatorie $X(t_1)=a\cos(2\pi f_0 t_1+\Theta)$ e $X(t_2)=a\cos(2\pi f_0 t_2+\Theta)$, le quali sono entrambe *trasformazioni* della stessa variabile aleatoria Θ. La funzione di autocorrelazione $R_X(t_1,t_2)$ si può calcolare facilmente con il teorema del valor medio:

$$R_X(t_1,t_2) = E\{X(t_1),X(t_2)\} = E_\Theta\{a\cos(2\pi f_0 t_1+\Theta)a\cos(2\pi f_0 t_2+\Theta)\}$$
$$= a^2\int_0^\pi \frac{1}{\pi}\cos(2\pi f_0 t_1+\theta)\cos(2\pi f_0 t_2+\theta)d\theta$$
$$= \frac{a^2}{2\pi}\int_0^\pi\left[\cos(2\pi f_0(t_1+t_2)+2\theta)+\cos(2\pi f_0(t_1-t_2))\right]d\theta$$
$$= \frac{a^2}{2}\cos(2\pi f_0(t_1-t_2)) \tag{E8.3.2}$$

La funzione $R_X(t_1,t_2)$ è cosinusoidale e dipende dalle variabili t_1 e t_2 attraverso

la loro differenza, ovvero

$$R_X(t_1,t_2) = R_X(t_1 - t_2) \tag{E8.3.3}$$

Questa proprietà, come vedremo nel paragrafo successivo, è verificata da una classe notevole di processi. ❑

8.3 Processi aleatori stazionari

8.3.1 Stazionarietà in senso stretto

Una proprietà notevole di alcuni processi aleatori è la *stazionarietà*. Come abbiamo già visto nei precedenti paragrafi, gli indici statistici di un processo, ad esempio la funzione valor medio $\eta_X(t)$ o la funzione di autocorrelazione $R_X(t_1,t_2)$ e, a maggior ragione, le funzioni densità di probabilità $f_X(x_1,x_2,...,x_n; t_1,t_2,...,t_n)$ dipendono in generale dalla scelta degli n istanti di tempo $t_1,t_2,...,t_n$ in corrispondenza dei quali viene valutato il processo. Cosa succede se *spostiamo rigidamente* tutti gli istanti temporali, cioè consideriamo la nuova n-upla di istanti $t_1 + \Delta t, t_2 + \Delta t, ,..., t_n + \Delta t$, con Δt arbitrario? In generale, otteniamo un diverso valore della funzione densità di probabilità di ordine n. Se, viceversa, il valore della funzione densità resta invariato *qualunque sia* Δt *e per ogni ordine* n, allora il processo si dice *stazionario in senso stretto:*

$$f_X(x_1,x_2,...,x_n;t_1+\Delta t,t_2+\Delta t,...,t_n+\Delta t) = f_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n) \quad \forall n, \forall \Delta t \tag{8.3.1}$$

La stazionarietà in senso stretto (o in senso *forte*) richiede dunque che le funzioni densità di probabilità del processo di qualunque ordine siano *invarianti* rispetto a una traslazione rigida degli istanti temporali. Detto in altre parole, ciò significa che i processi $X(t)$ e $X(t+\Delta t)$ *hanno le stesse statistiche*, e quindi sono *equivalenti* dal punto di vista statistico. Ciò non significa che $X(t+\Delta t)$ sia *uguale* a $X(t)$; infatti, le funzioni campione di $X(t+\Delta t)$ sono ottenute da quelle di $X(t)$ per traslazione temporale, e quindi i due processi sono differenti. Se però $X(t)$ è stazionario in senso stretto, non è possibile distinguerlo da $X(t+\Delta t)$ con misure *statistiche*.

Discutiamo ora le conseguenze di questa definizione. Se consideriamo la densità del primo ordine del processo, la definizione (8.3.1) ci dice che deve valere l'uguaglianza

$$f_X(x;t) = f_X(x;t+\Delta t) \quad \forall \Delta t \tag{8.3.2}$$

Poiché Δt è arbitrario, se ne conclude che la densità di probabilità del primo ordine $f_X(x;t)$ *non* dipende dal tempo: $f_X(x;t) \equiv f_X(x)$ (stazionarietà di *ordine uno*). Questa osservazione porta anche a concludere che tutte le grandezze statistiche del primo ordine del processo non dipendono a loro volta dalla variabile tempo: tutte le variabili aleatorie estratte da $X(t)$ sono equidistribuite e hanno, in particolare, uguale valor medio e uguali potenza media statistica e varianza:

$$\eta_X(t) \equiv \eta_X, \; P_X(t) \equiv P_X, \; \sigma_X^2(t) \equiv \sigma_X^2 \tag{8.3.3}$$

Esaminiamo ora le implicazioni della stazionarietà sulle statistiche del *secondo* ordine. Immaginiamo, allora, di osservare il processo aleatorio $X(t)$ a due istanti arbitrari t_1 e t_2, estraendo da esso le variabili aleatorie $X(t_1)$ ed $X(t_2)$. Come sappiamo, il comportamento statistico congiunto di tali variabili aleatorie è descritto dalla densità di probabilità del secondo ordine $f_X(x_1,x_2;t_1,t_2)$. Se consideriamo, poi, la coppia di variabili aleatorie $X(t_1+\Delta t)$ e $X(t_2+\Delta t)$ estratte agli istanti $t_1+\Delta t$ e $t_2+\Delta t$, sappiamo che

$$f_X(x_1,x_2;t_1,t_2) = f_X(x_1,x_2;t_1+\Delta t,t_2+\Delta t) \tag{8.3.4}$$

si ha cioè in particolare la *stazionarietà di ordine due*. Considerando che anche in questo caso Δt è arbitrario, è chiaro che la funzione $f_X(x_1,x_2;t_1,t_2)$ non può dipendere da t_1 *e* t_2 *separatamente,* ma deve dipendere soltanto dalla *differenza* t_1-t_2 tra gli istanti temporali, che resta appunto invariata in una traslazione rigida dei tempi:

$$f_X(x_1,x_2;t_1,t_2) = f_X(x_1,x_2;t_1-t_2) \tag{8.3.5}$$

Tutte le statistiche del *secondo ordine*, e in particolare la funzione di autocorrelazione e la funzione di autocovarianza godono evidentemente di questa stessa proprietà:

$$R_X(t_1,t_2) = R_X(t_1-t_2) \quad , \quad C_X(t_1,t_2) = C_X(t_1-t_2) \tag{8.3.6}$$

Generalizzando, è chiaro che la densità di probabilità (e tutte le statistiche) di ordine n di un processo stazionario in senso stretto dipendono soltanto dalle $(n-1)$ differenze (distanze) $t_1-t_2, t_3-t_3, \ldots, t_{n-1}-t_n$ tra gli istanti $t_1,t_2,\ldots,t_n$, differenze che restano *invariate* in una traslazione rigida dei tempi:

$$f_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n) = f_X(x_1,x_2,...,x_n;t_1-t_2,t_3-t_3,\ldots,t_{n-1}-t_n) \qquad (8.3.7)$$

Attraverso le "regole marginali" (7.4.36) non è difficile verificare che la stazionarietà di ordine n implica tutte le stazionarietà di ordine inferiore, mentre non vale il viceversa.

Esempio 8.4
Consideriamo il processo aleatorio parametrico

$$X(t) = A \qquad (E8.4.1)$$

ove A è una variabile aleatoria uniforme nell'intervallo $[-1,1]$ e cerchiamo di verificare se il processo è stazionario in senso stretto. Alcune funzioni campione di questo processo sono rappresentate in Figura 8.10: esse sono segnali *costanti* il cui valore è rappresentato dal particolare valore assunto in ogni prova dell'esperimento della variabile aleatoria A.

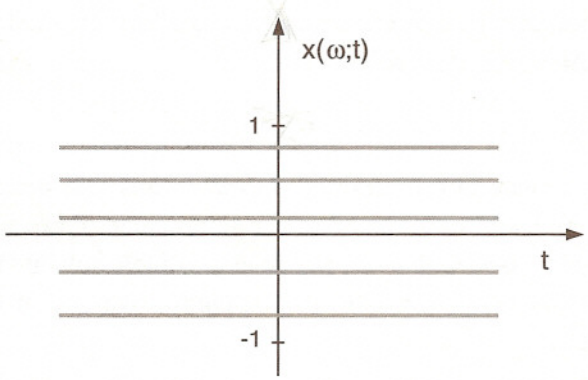


Figura 8.10 Funzioni campione del processo dell'Esempio 8.4

Evidentemente, il processo è stazionario di *ordine uno* perché, comunque si fissi un istante di tempo, la densità di probabilità della variabile aleatoria risultante è sempre la stessa:

$$f_X(x;t) = f_A(x) \qquad (E8.4.2)$$

in quanto ovviamente $X(t) = A$. Ragionando in questa stessa maniera si capisce che il processo è stazionario di ordine n *qualunque*. Infatti, comunque si fissi la n-upla di istanti $t_1,t_2,...,t_n$ si ottiene sempre la *stessa* n-upla di variabili aleatorie $X(t_1) = A, X(t_2) = A,\ldots, X(t_n) = A$, e quindi la densità di ordine n non dipende dal

tempo. A maggior ragione, essa non dipenderà da *traslazioni temporali*. D'altronde, guardando le funzioni campione di Figura 8.10, è chiaro che il processo $X(t)$ è invariante rispetto a una traslazione dell'asse dei tempi di Δt, perciò $X(t+\Delta t)$ ha le stesse statistiche di qualunque ordine di $X(t)$. ❑

8.3.2 Stazionarietà in senso lato

La verifica della stazionarietà in senso stretto di un processo è, in generale, estremamente difficoltosa (salvo casi particolari come i *processi Gaussiani* studiati nel Paragrafo 8.5). Di solito nelle applicazioni si considera una definizione di stazionarietà molto meno restrittiva e difficile da verificare della precedente: la *stazionarietà in senso lato*.

Un processo aleatorio $X(t)$ è *stazionario in senso lato* (o in *senso debole*) se la sua funzione valor medio $\eta_X(t)$ è *costante* e la sua funzione di autocorrelazione $R_X(t_1,t_2)$ non dipende da t_1 e t_2 separatamente, ma solo dalla differenza (t_1-t_2):

$$\eta_X(t)=\eta_X \tag{8.3.8a}$$

$$R_X(t_1,t_2)=R_X(t_1-t_2) \tag{8.3.8b}$$

La definizione di stazionarietà in senso lato coinvolge solo due *particolari* statistiche semplificate, una del primo e l'altra del secondo ordine, e *non* richiede alcuna proprietà di invarianza delle densità di probabilità; è quindi una proprietà molto meno "impegnativa" della stazionarietà in senso stretto. Infatti, un processo stazionario in senso stretto è anche stazionario in senso lato, mentre la stazionarietà in senso lato non implica quella in senso stretto.

Se il processo è stazionario in senso lato, la funzione di autocovarianza vale

$$C_X(t_1,t_2)=E\{[X(t_1)-\eta_X][X(t_2)-\eta_X]\}=R_X(t_1-t_2)-\eta_X{}^2=C_X(t_1-t_2) \tag{8.3.9}$$

cioè dipende anch'essa dalla sola *differenza* temporale t_1-t_2.

Anche per le grandezze statistiche semplificate si pone comunque il problema della *misurazione:* come si può riuscire in pratica ad avere un'idea dell'andamento della funzione valor medio e della funzione di autocorrelazione di un dato processo osservato? Questa domanda nasce dalla constatazione che sembra problematico se non impossibile effettuare una misura di media statistica che coinvolge la conoscenza di *tutte* le funzioni campione del processo! Saremo in grado di rispondere a questa domanda fondamentale solo dopo lo studio dei processi

ergodici del Paragrafo 8.7.

Esempio 8.5

Riprendiamo il processo parametrico $X(t)$ degli Esempi 8.2-8.3, ovvero

$$X(t) = a\cos(2\pi f_0 t + \Theta) \tag{E8.5.1}$$

ove a ed f_0 sono noti e $\Theta \in \mathcal{U}(0,\pi)$. Tale processo ha valor medio (si veda la (E8.2.3))

$$\eta_X(t) = -\frac{2a}{\pi}\sin(2\pi f_0 t) \tag{E8.5.2}$$

e quindi non è stazionario in senso lato. Consideriamo però lo stesso processo in cui stavolta $\Theta \in \mathcal{U}(0,2\pi)$. Ripetendo i conti dell'Esempio 8.2 si ottiene immediatamente:

$$E\{X(t)\} = \frac{a}{2\pi}\int_0^{2\pi}\cos(2\pi f_0 t + \theta)d\theta = \frac{a}{2\pi}\left[\sin(2\pi f_0 t + 2\pi) - \sin(2\pi f_0 t)\right] = 0 \tag{E8.5.3}$$

e, modificando i conti dell'Esempio 8.3:

$$R_X(t_1,t_2) = \frac{a^2}{4\pi}\int_0^{2\pi}\left[\cos(2\pi f_0(t_1+t_2)+2\theta) + \cos(2\pi f_0(t_1-t_2))\right]d\theta$$

$$= \frac{a^2}{2}\cos(2\pi f_0(t_1-t_2)) = R_X(t_1-t_2) \tag{E8.5.4}$$

Quindi la funzione valor medio è *costante*, e la funzione di autocorrelazione dipende solo dalla *differenza* $t_1 - t_2$: con la nuova scelta della distribuzione della variabile Θ il processo è *stazionario in senso lato*. ❑

Esempio 8.6

Consideriamo il processo aleatorio parametrico

$$X(t) = \sum_{k=-\infty}^{+\infty} P(t-2nT) \tag{E8.6.1}$$

ottenuto attraverso la *periodicizzazione* del segnale aleatorio

BIBLIOTECA INGEGNERIA PISA

$$P(t)=\left(1-\frac{|t|}{\Theta}\right)\text{rect}\left(\frac{t}{2\Theta}\right) \tag{E8.6.2}$$

ove Θ è una variabile aleatoria uniformemente distribuita nell'intervallo $[0,T]$, e verifichiamo se $X(t)$ è stazionario in senso lato o meno.

Per studiare la stazionarietà del processo $X(t)$ è utile rappresentare alcune sue funzioni campione, ciascuna delle quali corrisponde a valori diversi della variabile aleatoria Θ. La rappresentazione è semplice se si considera che ogni funzione è ottenuta dalla periodicizzazione con periodo $2T$ di un *impulso triangolare* $P(t)$ di ampiezza di picco unitaria e durata *aleatoria* 2Θ (Figura 8.11). Cominciamo col cercare di capire se la funzione valor medio del processo è costante al variare del tempo.

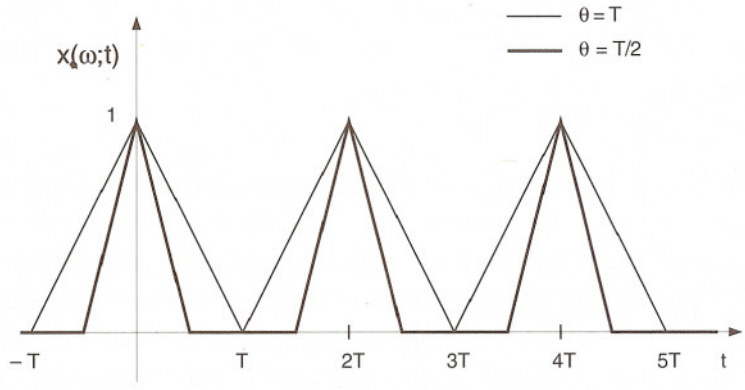


Figura 8.11 Alcune funzioni campione del processo dell'Esempio 8.6

Come si vede dalla Figura 8.11, *tutte* le funzioni campione valgono 1 negli istanti $2kT$ con k intero, e valgono 0 per tutti gli istanti $2kT+T$. Quindi,

$$E\{X(2kT)\}=1 \tag{E8.6.3}$$

$$E\{X(2kT+T)\}=0 \tag{E8.6.4}$$

La funzione valor medio $\eta_X(t)$ di $X(t)$ *non* è costante e il processo stesso *non è stazionario in senso lato*. ❑

Esempio 8.7 (Segnale dati)

Consideriamo un segnale $V(t)$ generato con il seguente procedimento: agli istanti $t_k=kT$, con $k=..,-1,0,1,..$, viene lanciata una moneta ideale (in maniera

indipendente dai lanci precedenti e successivi) e, nel caso di un risultato "testa", si fa assumere al segnale il valore +1 che viene *mantenuto* per i T secondi successivi a t_k; viceversa, se il risultato è "croce" viene mantenuto il valore –1.

Indichiamo, innanzitutto, con V_n il valore del segnale nell'intervallo di ampiezza T che segue l'n-esimo lancio della moneta:

$$V(t) = V_n \quad , \quad nT \le t < (n+1)T \tag{E8.7.1}$$

Evidentemente, V_n è una variabile aleatoria che assume i valori +1 e –1 con probabilità 1/2, ed è indipendente da V_k, $k \neq n$. Osservando che il valore del segnale viene mantenuto per tutto un intervallo di ampiezza T, e richiamando l'espressione del segnale Sample & Hold (5.4.9)-(5.4.10), si trova immediatamente la forma della generica funzione campione (e quindi del processo in forma parametrica):

$$V(t) = \sum_{n=-\infty}^{+\infty} V_n \operatorname{rect}\left(\frac{t - nT - T/2}{T}\right) \tag{E8.7.2}$$

L'andamento di una funzione campione del processo aleatorio è illustrato nella Figura 8.12.

Un segnale di questo tipo è un buon modello del *segnale dati binario* con velocità di *clock* pari a $1/T$ che si può trovare ad esempio nei collegamenti tra un Personal Computer (PC) e una stampante, su uno dei fili in uscita alla porta seriale del PC. Il livello della tensione rappresenta il valore del segnale *binario* che viene scambiato tra il PC e la stampante, e la cadenza $1/T$ è la "velocità" alla quale funziona la porta del calcolatore. Poiché in generale non è nota la successione di dati che transitano sulla linea seriale, si ammette che i valori del segnale siano *aleatori*, e quindi il segnale $V(t)$ è un processo aleatorio.

Di questo processo, ci proponiamo di determinare la densità di probabilità del primo ordine $f_V(v;t)$, la funzione valor medio $\eta_V(t)$ e la funzione di autocorrelazione $R_V(t_1,t_2)$.

Per determinare la funzione densità di probabilità del primo ordine $f_V(v;t)$ si osserva il processo all'istante t estraendo da esso la variabile aleatoria $V(t)$. Se $t \in [nT,(n+1)T)$, la variabile aleatoria $V(t)$ coincide con V_n (si veda la (E8.7.1)) e, quindi, per le proprietà di V_n abbiamo:

$$f_V(v;t) = \frac{1}{2}\delta(v+1) + \frac{1}{2}\delta(v-1) \tag{E8.7.3}$$

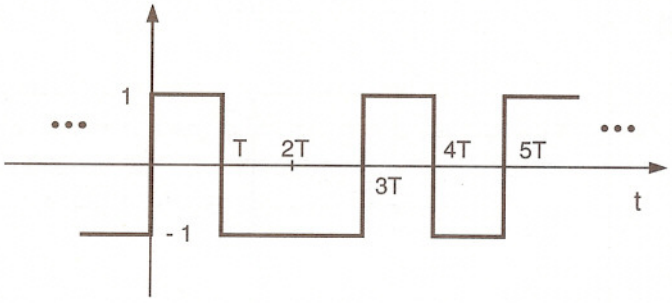


Figura 8.12 Funzione campione del processo *segnale dati binario* nell'Esempio 8.7

Questa funzione densità, ricavata per $t \in [nT,(n+1)T)$, è in realtà *indipendente* dal tempo: *il processo aleatorio è stazionario del primo ordine*. La funzione valor medio risulta quindi costante, e vale:

$$\eta_V = \int_{-\infty}^{+\infty} v\, f_V(v;t)\,dv = \int_{-\infty}^{+\infty} v\left[\frac{1}{2}\delta(v+1)+\frac{1}{2}\delta(v-1)\right]dv = -\frac{1}{2}+\frac{1}{2} = 0 \qquad \text{(E8.7.4)}$$

Il calcolo della funzione di autocorrelazione è invece un poco più complicato. Partiamo dalla definizione:

$$R_V(t_1,t_2) = E\{V(t_1)V(t_2)\} \qquad \text{(E8.7.5)}$$

Senza calcolare esplicitamente questa funzione, ci si può accorgere facilmente che il processo *non* è stazionario in senso lato. Supponiamo infatti che gli istanti di osservazione t_1 e t_2 appartengano *entrambi* all'intervallo $(kT,(k+1)T]$, per un opportuno valore di k (Figura 8.13a):

$$kT < t_1, t_2 < (k+1)T \qquad \text{(E8.7.6)}$$

Allora è chiaro che $V(t_1) = V(t_2) = V_k$, e quindi

$$R_V(t_1,t_2) = E\{V_k^2\} = 1 \qquad \text{(E8.7.7)}$$

in quanto $V_k^2 \equiv 1$. Trasliamo ora *rigidamente* i due istanti (mantenendo cioè la distanza tra i due) in modo da piazzarli a "cavallo" di due intervalli di clock consecutivi (Figura 8.13b):

$$kT < t_1 < (k+1)T \quad , \quad (k+1)T < t_2 < (k+2)T \qquad \text{(E8.7.8)}$$

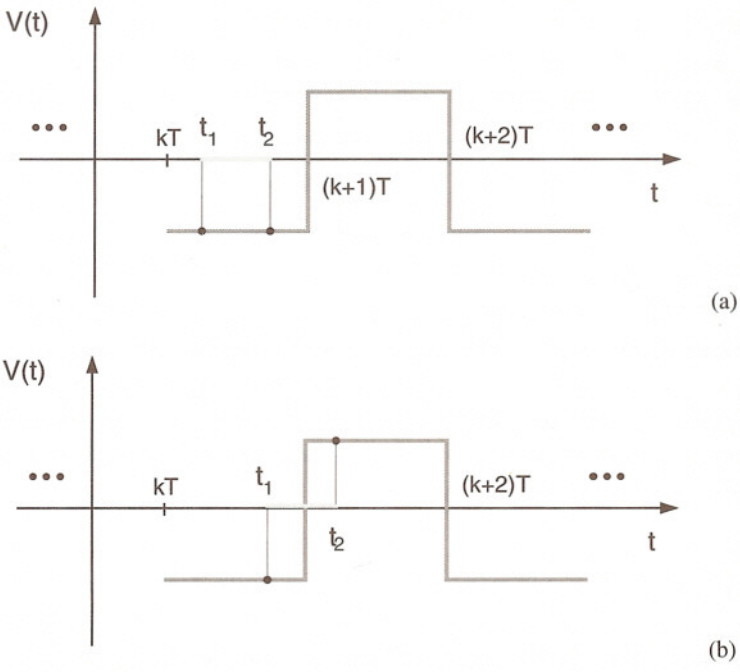


Figura 8.13 Dimostrazione che il processo dell'Esempio 8.7 non è stazionario

Se il processo fosse stazionario in senso lato, il valore della funzione di autocorrelazione *non dovrebbe cambiare* perché la *differenza* tra i due istanti non è cambiata. La funzione di autocorrelazione è data però in questo secondo caso da:

$$R_V(t_1,t_2) = E\{V_k V_{k+1}\} = E\{V_k\}E\{V_{k+1}\} = 0 \qquad \text{(E8.7.9)}$$

che è un valore *diverso* da quello del primo caso. Il processo non è quindi stazionario in autocorrelazione e non è stazionario in senso lato. Questo risultato può essere confermato con il calcolo:

$$\begin{aligned} R_V(t_1,t_2) &= E\{V(t_1)V(t_2)\} \\ &= E\left\{\sum_{n=-\infty}^{+\infty} V_n \,\text{rect}\left(\frac{t_1 - nT - T/2}{T}\right)\sum_{k=-\infty}^{+\infty} V_k \,\text{rect}\left(\frac{t_2 - kT - T/2}{T}\right)\right\} \\ &= \sum_{n=-\infty}^{+\infty}\sum_{k=-\infty}^{+\infty} E\{V_n V_k\}\,\text{rect}\left(\frac{t_1 - nT - T/2}{T}\right)\text{rect}\left(\frac{t_2 - kT - T/2}{T}\right) \end{aligned}$$

$$= \sum_{n=-\infty}^{+\infty} \text{rect}\left(\frac{t_1 - nT - T/2}{T}\right)\text{rect}\left(\frac{t_2 - nT - T/2}{T}\right) \qquad \text{(E8.7.10)}$$

Nell'ultimo passaggio, dopo avere sfruttato la linearità dell'operatore valor medio, abbiamo tenuto conto che $E\{V_n V_k\}$ è diverso da zero solo se $k = n$. Si vede che la funzione $R_V(t_1, t_2)$ è una funzione periodica in t_1 e t_2 del periodo di clock T, ma dipende dai due istanti *separatamente*.

❑

8.3.3 Proprietà della funzione di autocorrelazione di un processo stazionario in senso lato

La funzione di autocorrelazione (8.2.6) può anche essere definita in una maniera alternativa, in cui si mette in evidenza la *differenza* (distanza) τ tra i due istanti di tempo t_1 e t_2 considerati. Se infatti si pone $t_1 = t$, $t_2 = t - \tau$, si ha:

$$R_X(t_1, t_2) = R_X(t, t-\tau) \triangleq E\{X(t)X(t-\tau)\} \qquad \text{(8.3.10)}$$

Se il processo $X(t)$ è stazionario almeno in senso lato, la funzione di autocorrelazione dipende soltanto dalla *differenza* τ tra i due istanti, e non dall'istante particolare t in corrispondenza del quale si "piazza" il primo:

$$R_X(t, t-\tau) = R_X(\tau) \qquad \text{(8.3.11)}$$

Esempio 8.8

Una situazione che capita spesso è quella in cui l'andamento di un segnale è noto, ma non se ne conosce esattamente la posizione rispetto a un riferimento temporale. In questi casi il modello appropriato è un processo aleatorio del tipo

$$X(t) = p(t - \Theta) \qquad \text{(E8.8.1)}$$

dove $p(t)$ è un segnale *determinato*, mentre Θ è una variabile aleatoria che modella l'incertezza temporale sulla posizione del segnale (si pensi al problema dell'*eco radar* dell'Esempio 3.8). Come caso particolare, supponiamo che $p(t)$ sia un segnale *determinato periodico* di periodo T_0:

$$p(t) = p(t + T_0) \qquad \text{(E8.8.2)}$$

e che $\Theta \in \mathcal{U}(0, T_0)$. Dimostriamo che $X(t)$ è *stazionario in senso lato*. La funzione valor medio del processo è:

$$\eta_X(t) = E\{X(t)\} = E_\Theta\{p(t-\Theta)\} = \int_0^{T_0} p(t-\theta)\frac{1}{T_0}d\theta = \frac{1}{T_0}\int_{t-T_0}^{t} p(\alpha)d\alpha$$

ove nell'ultimo passaggio abbiamo effettuato il cambio di variabile $\alpha = t-\theta$. Osserviamo adesso che la funzione integranda è periodica di periodo T_0, e che l'intervallo di integrazione ha proprio ampiezza pari a un periodo. Segue che il risultato dell'integrale è *indipendente dalla posizione degli estremi di integrazione*, e quindi non dipende da t: il processo è *stazionario in valor medio.* Scegliendo per comodità $t = T_0/2$ si ha:

$$\eta_X(t) \equiv \eta_X = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} p(\alpha)d\alpha \qquad \text{(E8.8.3)}$$

cioè il valor medio *statistico* del processo $X(t)$ coincide con il valor medio *temporale* p_m (2.2.3) del segnale periodico determinato $p(t)$.

Per la funzione di autocorrelazione abbiamo:

$$R_X(t_1,t_2) = \mathrm{E}\{X(t_1)X(t_2)\} = E_\Theta\{p(t_1-\Theta)p(t_2-\Theta)\}$$
$$= \frac{1}{T_0}\int_0^{T_0} p(t_1-\theta)p(t_2-\theta)d\theta = \frac{1}{T_0}\int_{t_1-T_0}^{t_1} p(\alpha)\,p(\alpha+t_2-t_1)d\alpha \qquad \text{(E8.8.4)}$$

con la sostituzione $\alpha = t_1-\theta$. Anche in questo caso la funzione integranda, che è il prodotto di due segnali periodici di periodo T_0, è periodica di periodo T_0. Il risultato non dipende dal piazzamento degli estremi di integrazione, per cui si può scegliere $t_1 = T_0/2$ e concludere che:

$$R_X(t_1,t_2) = R_X(t_1-t_2) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} p(\alpha)\,p(\alpha+t_2-t_1)d\alpha \qquad \text{(E8.8.5)}$$

Quindi la funzione di autocorrelazione dipende dalla differenza t_1-t_2 e il processo $X(t)$ è *stazionario in senso lato.* Ponendo come di consueto $\tau = t_1-t_2$, si ha:

$$R_X(t_1,t_2) = R_X(\tau) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} p(\alpha)\,p(\alpha-\tau)d\alpha \qquad \text{(E8.8.6)}$$

e cioè la funzione di autocorrelazione *statistica* del processo *stazionario* $X(t)$ coincide con la funzione di autocorrelazione (4.4.33) del segnale periodico determinato $p(t)$. ❑

La funzione di autocorrelazione $R_X(\tau)$ è una grandezza di fondamentale importanza nello studio dei processi stazionari. Prima di discutere la ragione di questa importanza, analizziamo alcune proprietà formali di tale funzione per un processo stazionario in senso lato.

- La funzione di autocorrelazione $R_X(\tau)$ è *pari*:

$$R_X(\tau) = R_X(-\tau) \tag{8.3.12}$$

Per dimostrare questa relazione, basta osservare che, per la stazionarietà del processo, la correlazione tra le variabili aleatorie $X(t)$ e $X(t-\tau)$ assume il medesimo valore della correlazione tra le variabili $X(t+\tau)$ e $X(t)$ ottenute per traslazione rigida della quantità τ:

$$R_X(\tau) = E\{X(t)X(t-\tau)\} = E\{X(t+\tau)X(t)\} = R_X(-\tau) \tag{8.3.13}$$

- Il valore assunto da $R_X(\tau)$ nell'origine uguaglia la potenza media statistica del processo, cioè

$$R_X(0) = E\{X(t)^2\} = P_X \geq 0 \tag{8.3.14}$$

- $R_X(\tau)$ è massima in modulo nell'origine:

$$R_X(0) \geq |R_X(\tau)| \tag{8.3.15}$$

Consideriamo infatti la disuguaglianza

$$E\{[X(t) \pm X(t-\tau)]^2\} \geq 0 \tag{8.3.16}$$

che è sempre verificata, comunque si scelgano t e τ. Sviluppando il primo membro si ha:

$$E\{[X(t) \pm X(t-\tau)]^2\} = E\{[X(t)]^2 + [X(t-\tau)]^2 \pm 2X(t)X(t-\tau)\}$$
$$= R_X(0) + R_X(0) \pm 2R_X(\tau) \tag{8.3.17}$$

e, sostituendo:

$$R_X(0) \geq \mp R_X(\tau) \tag{8.3.18}$$

dalla quale segue immediatamente la (8.3.15).

- Se $R_X(\tau)$ non contiene componenti periodiche, il valore limite di $R_X(\tau)$ per $\tau \to \infty$ è pari al quadrato del valore medio:

$$\lim_{\tau\to\infty} R_X(\tau) = \eta_X^2 \tag{8.3.19}$$

Per giustificare questa proprietà riscriviamo la relazione (8.3.9) nella forma

$$R_X(\tau) = C_X(\tau) + \eta_X^2 \tag{8.3.20}$$

Al crescere di τ, la distanza tra gli istanti t e $t-\tau$ aumenta e quindi le funzioni campione del processo hanno "tempo" per variare sensibilmente. Questo comporta che i valori delle variabili aleatorie $X(t)$ e $X(t-\tau)$ tendono a diventare *incorrelati*, cioè la loro covarianza $C_X(\tau)$ si riduce. Al limite, quando $\tau \to +\infty$ la covarianza si *annulla* e la funzione di autocorrelazione tende a coincidere con il quadrato del valor medio.

Queste proprietà, come già accennato nel Paragrafo 4.4, non sono specifiche dei processi, ma sono comuni alle funzioni di autocorrelazione dei segnali determinati e aleatori, come si può facilmente dimostrare (si veda anche l'Esercizio proposto 4.2).

Esempio 8.9

Riprendiamo in considerazione il *segnale dati binario* $V(t)$ dell'Esempio 8.7. Come abbiamo già discusso, spesso il riferimento temporale assoluto di un segnale non è noto, in questo caso perché il segnale di clock con il quale vengono prodotti i dati ha una fase iniziale incognita. Allora il modello più appropriato per il segnale dati, che stavolta chiameremo $X(t)$, è:

$$X(t) = V(t-\Theta) = \sum_{n=-\infty}^{+\infty} V_n \,\mathrm{rect}\left(\frac{t-\Theta-nT-T/2}{T}\right)$$

ove Θ è la consueta variabile aleatoria uniformemente distribuita sull'intervallo $[0,T]$, e indipendente dai valori dei dati binari $\{V_n\}$. Il calcolo del valor medio del processo non differisce da quello dell'Esempio 8.7, per cui si conclude:

$$\eta_X(t) = \eta_X = 0 \tag{E8.9.1}$$

Anche il calcolo della funzione di autocorrelazione ricalca sostanzialmente quello dell'Esempio 8.7, con la differenza che stavolta si deve compiere un'operazione di media *in più* rispetto alla variabile "ritardo aleatorio" Θ:

$$R_X(t_1,t_2) = E_\Theta\left\{\sum_{n=-\infty}^{+\infty} \mathrm{rect}\left(\frac{t_1-\Theta-nT-T/2}{T}\right)\mathrm{rect}\left(\frac{t_2-\Theta-nT-T/2}{T}\right)\right\}$$

$$= \sum_{n=-\infty}^{+\infty} E_\Theta \left\{ \text{rect}\left(\frac{t-\Theta-nT-T/2}{T}\right)\text{rect}\left(\frac{t-\tau-\Theta-nT-T/2}{T}\right)\right\}$$

$$= \sum_{n=-\infty}^{+\infty} \frac{1}{T}\int_0^T \text{rect}\left(\frac{t-\theta-nT-T/2}{T}\right)\text{rect}\left(\frac{t-\tau-\theta-nT-T/2}{T}\right)d\theta$$

$$= \frac{1}{T}\sum_{n=-\infty}^{+\infty} \int_{t-nT-T}^{t-nT} \text{rect}\left(\frac{\alpha-T/2}{T}\right)\text{rect}\left(\frac{\alpha-\tau-T/2}{T}\right)d\theta \qquad \text{(E8.9.2)}$$

ove nell'ultimo passaggio si è posto $\alpha = t-\theta-nT$. La funzione integranda non contiene la variabile n e, al variare di n, si calcolano integrali della stessa funzione integranda su intervalli di ampiezza T disgiunti che ricoprono tutto l'asse reale di tempi (si veda a questo proposito l'analogo caso rappresentato nella Figura 3.46). La conclusione è che

$$R_X(t,t-\tau) = R_X(\tau) = \frac{1}{T}\int_{-\infty}^{\infty} \text{rect}\left(\frac{\alpha-T/2}{T}\right)\text{rect}\left(\frac{\alpha-\tau-T/2}{T}\right)d\theta \qquad \text{(E8.9.3)}$$

cioè il processo è *stazionario in senso lato* perché la dipendenza dal tempo t è scomparsa. Si vede inoltre che la funzione di autocorrelazione del *processo* è proporzionale alla funzione di autocorrelazione (4.4.23) del segnale *determinato* $\text{rect}((t-T/2)/T)$:

$$R_X(\tau) = \frac{1}{T}\text{rect}\left(\frac{\tau-T/2}{T}\right) \otimes \text{rect}\left(\frac{-\tau-T/2}{T}\right) = \left(1-\frac{|\tau|}{T}\right)\text{rect}\left(\frac{\tau}{2T}\right) \qquad \text{(E8.9.4)}$$

Questa funzione è l'impulso *triangolare* di ampiezza unitaria di Figura 8.14. Il lettore verifichi che questa funzione soddisfa le quattro proprietà generali delle funzioni di autocorrelazione dei processi stazionari.

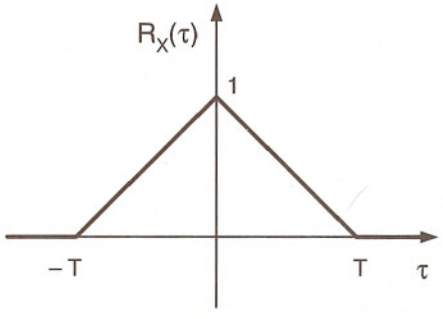


Figura 8.14 Funzione di autocorrelazione del segnale dati stazionario dell'Esempio 8.9

❑

Qual è il significato e l'utilità della funzione di autocorrelazione $R_X(\tau)$ di un processo stazionario? Per chiarire questo punto partiamo da un esempio. I due insiemi di funzioni campione di Figura 8.15a e 8.15b sono relativi a due diversi processi stazionari, rispettivamente $X(t)$ e $Y(t)$, aventi *stesse statistiche del primo ordine* (valor medio, potenza, densità del primo ordine).

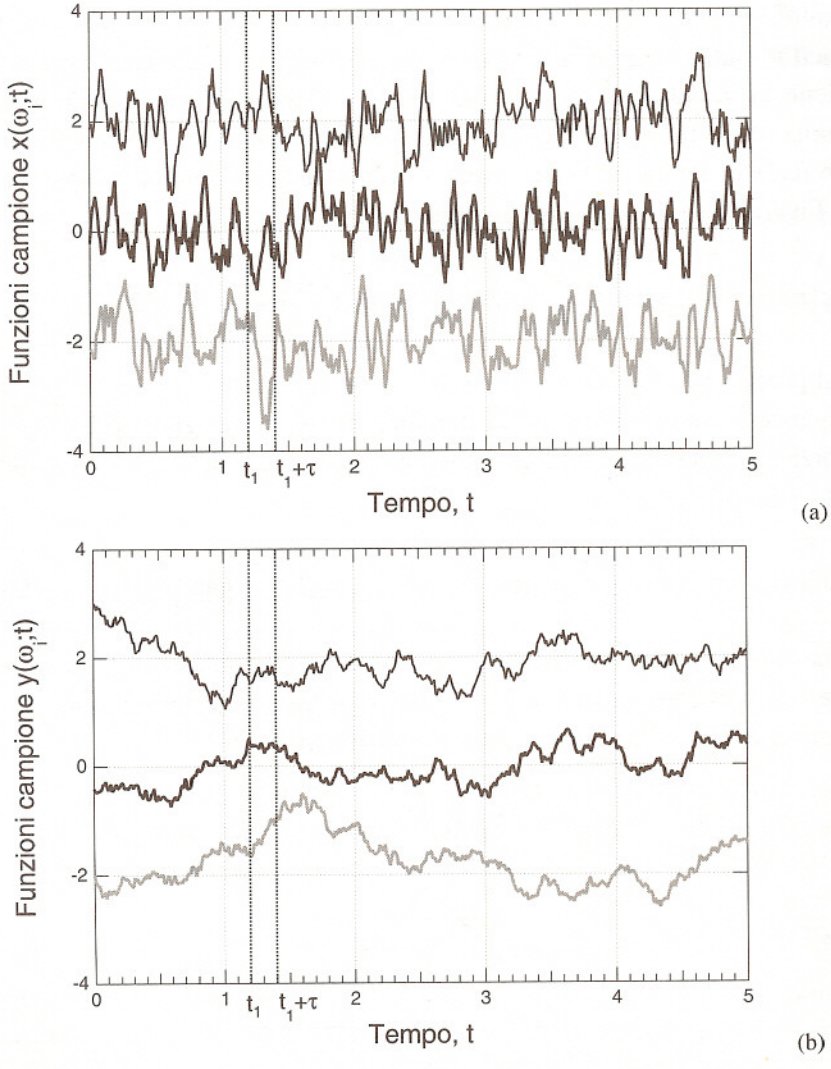


Figura 8.15 Funzioni campione di due processi con diverse funzioni di autocorrelazione

Evidentemente, però, i due processi differiscono parecchio nella rispettiva *velocità media di variazione* delle funzioni campione. Se fissiamo sul processo $Y(t)$

i due istanti alla distanza τ indicati nella Figura 8.15b, notiamo che le funzioni campione, piuttosto lente, hanno avuto poco tempo per variare, e quindi i valori delle variabili aleatorie estratte dal processo a questi istanti sono *molto correlati*. Viceversa, nello stesso lasso di tempo τ le funzioni campione del processo $X(t)$, assai più veloci, sono variate considerevolmente, e i valori delle due variabili aleatorie sono molto meno correlati. In conclusione, la funzione $R_X(\tau)$ *decresce velocemente* a zero quando τ aumenta, come illustrato nella Figura 8.16a, mentre la funzione $R_Y(\tau)$ decresce più lentamente (Figura 8.16b).

Dunque la funzione di autocorrelazione misura la *rapidità di variazione* del segnale aleatorio. Per quantificare con un singolo parametro la "velocità" del segnale si introduce il *tempo di correlazione* τ_{cor}, definito come la minima distanza che deve intercorrere tra due istanti di osservazione affinché le variabili aleatorie estratte dal processo siano incorrelate. Evidentemente, il tempo di correlazione è pari alla *semidurata* della funzione di autocorrelazione (Figura 8.16), eventualmente definita in modo convenzionale quando l'autocorrelazione non ha durata rigorosamente limitata (si veda il Paragrafo 4.3.3). A tempo di correlazione *grande* corrispondono funzioni campione che variano lentamente, e, viceversa, a tempo di correlazione *piccolo*, corrisponde un processo le cui funzioni campione manifestano variazioni veloci.

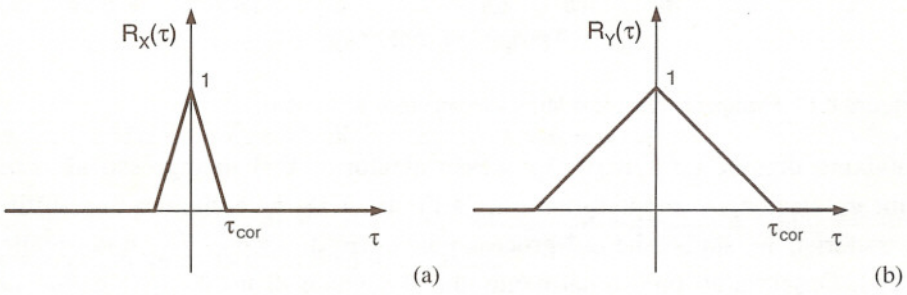


Figura 8.16 Funzioni di autocorrelazione dei segnali aleatori di Figura 8.15

8.4 Filtraggio di un segnale aleatorio

8.4.1 Relazione ingresso-uscita tra le statistiche semplificate

Un caso tipico dell'elaborazione dei segnali è quello in cui l'osservato $X(t)$ è costituito da una componente determinata $s(t)$ (il segnale "utile") accompagnata da un *disturbo* aleatorio a valor medio nullo $D(t)$ (chiamato anche *rumore*):

$$X(t) = s(t) + D(t) \tag{8.4.1}$$

come nell'onda quadra di ampiezza unitaria "rumorosa" in Figura 8.17. Naturalmente, si cercherà di elaborare $X(t)$ in modo da preservare la componente utile $s(t)$ e reiettare il più possibile il disturbo $D(t)$. Quest'operazione può essere effettuata da un *filtro*, cioè da un sistema lineare stazionario nel senso del Paragrafo 4.3, il cui comportamento riguardo ai segnali determinati è perfettamente noto. Resta dunque da studiare la questione del *filtraggio di un segnale aleatorio*, che è l'oggetto di questo paragrafo.

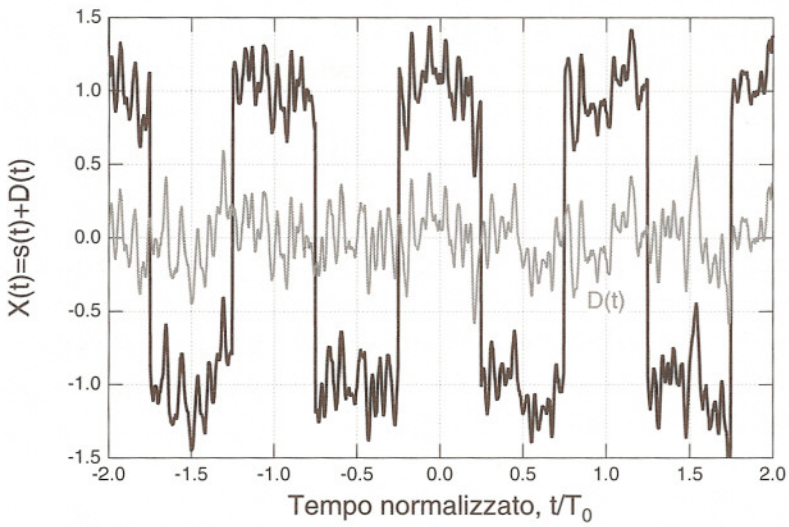


Figura 8.17 Esempio di segnale determinato con disturbo aleatorio

Inviamo dunque un generico processo aleatorio $X(t)$ in ingresso al sistema lineare stazionario rappresentato nella Figura 8.18, e cerchiamo di stabilire le caratteristiche statistiche del processo aleatorio di uscita $Y(t)$, note quelle di $X(t)$. Osserviamo preliminarmente che il segnale di uscita $Y(t)$ è un *nuovo* processo le cui funzioni campione possono essere facilmente messe in corrispondenza con i risultati dell'esperimento che ha generato $X(t)$. Infatti, data una funzione campione $x(\overline{\omega};t)$, scelta in modo arbitrario fra tutte quelle del processo $X(t)$, la corrispondente funzione campione $y(\overline{\omega};t)$ di $Y(t)$ è ricavabile facilmente (si veda la (4.2.6)):

$$y(\overline{\omega};t) = x(\overline{\omega};t;) \otimes h(t) \tag{8.4.2}$$

dove $h(t)$ è la risposta impulsiva del sistema in esame. Questo risultato vale *per qualunque realizzazione* $x(\overline{\omega};t)$ *del processo* $X(t)$, e quindi scriveremo, per

riassumere questa osservazione:

$$Y(t) = X(t) \otimes h(t) \tag{8.4.3}$$

Questa relazione sembra *mal posta*, perché non abbiamo definito esplicitamente una convoluzione per segnali aleatori; essa però va intesa nel senso appena discusso, cioè che il legame vale in realtà per la coppia di funzioni campione *determinate* $x(\overline{\omega};t)$ e $y(\overline{\omega};t)$.

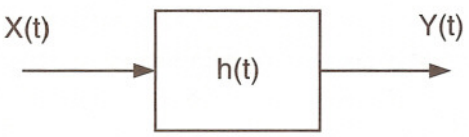


Figura 8.18 Filtraggio di un processo aleatorio

Purtroppo il problema di ricavare le densità di probabilità del processo di uscita $f_Y(y_1, y_2, ..., y_n; t_1, t_2, ..., t_n)$ a partire da quelle $f_X(x_1, x_2, ..., x_n; t_1, t_2, ..., t_n)$ del processo d'ingresso è, salvo casi particolari, *insolubile*. Cerchiamo però più modestamente di ricavare la funzione valor medio e la funzione di autocorrelazione del processo $Y(t)$, supponendo di conoscere le stesse statistiche di $X(t)$, ovvero le funzioni $\eta_X(t)$ ed $R_X(t_1, t_2)$.

Il valor medio $\eta_Y(t)$ di $Y(t)$ è:

$$\eta_Y(t) = E\{Y(t)\} = E\{X(t) \otimes h(t)\} = E\left\{\int_{-\infty}^{+\infty} h(\alpha) X(t-\alpha) d\alpha\right\} \tag{8.4.4}$$

L'ordine delle operazioni di media statistica e di integrazione può essere invertito per la linearità degli operatori stessi:

$$\eta_Y(t) = \int_{-\infty}^{+\infty} E\{h(\alpha) X(t-\alpha)\} d\alpha \tag{8.4.5}$$

In questa relazione, l'operazione di valor medio agisce solo sul segnale *aleatorio* $X(t)$ e non sul segnale *determinato* $h(t)$. Possiamo allora scrivere:

$$\eta_Y(t) = \int_{-\infty}^{+\infty} h(\alpha) E\{X(t-\alpha)\}\, d\alpha = \int_{-\infty}^{+\infty} h(\alpha)\, \eta_X(t-\alpha) d\alpha = \eta_X(t) \otimes h(t) \tag{8.4.6}$$

Dunque, la funzione valor medio del processo aleatorio di uscita è pari alla *convoluzione* della funzione valor medio del processo in ingresso con la risposta

impulsiva del sistema. Interpretiamo questa relazione: se definiamo il processo *a media nulla* $X_0(t) = X(t) - \eta_X(t)$, possiamo pensare il processo d'ingresso $X(t)$ scomposto come la somma di una componente *determinata* pari alla funzione valor medio (nota) e una componente puramente aleatoria a valor medio nullo, appunto $X_0(t)$:

$$X(t) = \eta_X(t) + X_0(t) \tag{8.4.7}$$

Allora, la componente determinata del processo viene ovviamente *filtrata* dal sistema come un qualunque segnale determinato, per dare la componente determinata (cioè la funzione valor medio) del processo d'uscita:

$$\eta_Y(t) = \eta_X(t) \otimes h(t) \tag{8.4.8}$$

Ricaviamo, adesso, la *funzione di autocorrelazione* $R_Y(t_1, t_2)$ di $Y(t)$

$$\begin{aligned} R_Y(t_1,t_2) &= E\{Y(t_1)Y(t_2)\} = E\{[X(t_1) \otimes h(t_1)] \cdot [X(t_2) \otimes h(t_2)]\} \\ &= E\left\{\int_{-\infty}^{+\infty} X(\alpha)h(t_1-\alpha)d\alpha \int_{-\infty}^{+\infty} X(\beta)h(t_2-\beta)d\beta\right\} \end{aligned} \tag{8.4.9}$$

Invertendo le operazioni di media statistica e di integrazione si ricava:

$$\begin{aligned} R_Y(t_1,t_2) &= \int_{\alpha=-\infty}^{+\infty} \int_{\beta=-\infty}^{+\infty} E\{X(\alpha)h(t_1-\alpha)X(\beta)h(t_2-\beta)\}d\alpha\, d\beta \\ &= \int_{\alpha=-\infty}^{+\infty} \int_{\beta=-\infty}^{+\infty} h(t_1-\alpha)h(t_2-\beta)E\{X(\alpha)X(\beta)\}d\alpha\, d\beta \\ &= \int_{\beta=-\infty}^{+\infty} \int_{\alpha=-\infty}^{+\infty} R_X(\alpha,\beta)h(t_1-\alpha)h(t_2-\beta)d\alpha\, d\beta \end{aligned} \tag{8.4.10}$$

Possiamo riscrivere questo integrale doppio nella forma di una *doppia convoluzione*:

$$R_Y(t_1,t_2) = \int_{-\infty}^{+\infty} [R_X(t_1,\beta) \otimes h(t_1)]h(t_2-\beta)d\beta = R_X(t_1,t_2) \otimes h(t_1) \otimes h(t_2) \tag{8.4.11}$$

La prima operazione di convoluzione coinvolge solo la variabile t_1 e, nello svolgimento di essa, la variabile t_2 viene considerata come una *costante*. Nello svolgimento del secondo prodotto di convoluzione, invece, i ruoli delle variabili t_1 e t_2 si scambiano.

8.4.2 Filtraggio di un processo aleatorio stazionario in senso lato

Esaminiamo ora il caso particolare in cui il processo d'ingresso al filtro è *stazionario in senso lato*. Per la funzione valor medio si ha:

$$\eta_X(t) \otimes h(t) = \int_{-\infty}^{+\infty} h(\alpha)\, \eta_X(t-\alpha)\, d\alpha = \eta_X \int_{-\infty}^{+\infty} h(\alpha)\, d\alpha = \eta_X\, H(0) \equiv \eta_Y \qquad (8.4.12)$$

Il processo di uscita $Y(t)$ ha, a sua volta, valor medio costante pari a $\eta_Y = \eta_X\, H(0)$, ove naturalmente $H(f)$ è la risposta in frequenza del sistema: in questo caso il processo d'ingresso contiene una "componente continua" che viene modificata in ragione del guadagno in continua del filtro $H(0)$.

La funzione di autocorrelazione $R_Y(t, t-\tau)$ è poi:

$$\begin{aligned}
R_Y(t,t-\tau) &= E\{Y(t)Y(t-\tau)\} = E\{[X(t)\otimes h(t)]\cdot[X(t-\tau)\otimes h(t-\tau)]\} \\
&= E\left\{\int_{-\infty}^{+\infty} h(\alpha)X(t-\alpha)\,d\alpha \cdot \int_{-\infty}^{+\infty} h(\beta)X(t-\tau-\beta)\,d\beta\right\} \\
&= E\left\{\int_{\alpha=-\infty}^{+\infty}\int_{\beta=-\infty}^{+\infty} h(\beta)h(\alpha)X(t-\alpha)X(t-\tau-\beta)\,d\beta\,d\alpha\right\} \\
&= \int_{\alpha=-\infty}^{+\infty}\int_{\beta=-\infty}^{+\infty} h(\beta)h(\alpha)E\{X(t-\alpha)X(t-\tau-\beta)\}\,d\beta\,d\alpha \\
&= \int_{\beta=-\infty}^{+\infty}\int_{\alpha=-\infty}^{+\infty} h(\alpha)R_X(\tau+\beta-\alpha)h(\beta)\,d\beta\,d\alpha \\
&= \int_{\beta=-\infty}^{+\infty} h(\beta)\int_{\alpha=-\infty}^{+\infty} h(\alpha)R_X(\tau+\beta-\alpha)\,d\alpha\,d\beta
\end{aligned} \qquad (8.4.13)$$

Già da quest'espressione si nota che la funzione di autocorrelazione del processo di uscita *non dipende dal tempo t* e quindi, tenendo conto della (8.4.12), il processo $Y(t)$ è *stazionario in senso lato*. Osservando ora che

$$\int_{\alpha=-\infty}^{+\infty} h(\alpha)R_X(\tau+\beta-\alpha)\,d\alpha = R_X(\tau+\beta)\otimes h(\tau+\beta) \qquad (8.4.14)$$

la (8.4.14) diventa:

$$R_Y(t,t-\tau) = R_Y(\tau) = \int_{-\infty}^{+\infty} h(\beta)\left[R_X(\tau+\beta)\otimes h(\tau+\beta)\right]d\beta \qquad (8.4.15)$$

Con il cambiamento di variabile $\gamma = -\beta$, si ottiene:

$$R_Y(\tau) = \int_{-\infty}^{+\infty} h(-\gamma)\left[R_X(\tau-\gamma) \otimes h(\tau-\gamma)\right] d\gamma = R_X(\tau) \otimes h(\tau) \otimes h(-\tau) \qquad (8.4.16)$$

Riassumendo, la risposta a un processo stazionario in senso lato di un sistema lineare stazionario è un processo anch'esso stazionario in senso lato, con funzione valor medio e funzione di autocorrelazione rispettivamente pari a:

$$\eta_Y = \eta_X H(0) \qquad (8.4.16a)$$

$$R_Y(\tau) = R_X(\tau) \otimes h(\tau) \otimes h(-\tau) = R_X(\tau) \otimes r_h(\tau) \qquad (8.4.16b)$$

ove $r_h(\tau)$ è la funzione di autocorrelazione della risposta impulsiva del filtro.

Esempio 8.10

Consideriamo il processo aleatorio

$$X(t) = \sum_{k=-\infty}^{+\infty} \delta(t - kT - \Theta) \qquad (E8.10.1)$$

ove Θ è una variabile aleatoria uniformemente distribuita in $[-T/2, T/2)$. Determiniamo la funzione valor medio $\eta_Y(t)$ del processo $Y(t)$ ottenuto dal *filtraggio* di $X(t)$ con un filtro la cui risposta impulsiva $h(t)$ è rappresentata nella Figura 8.19.

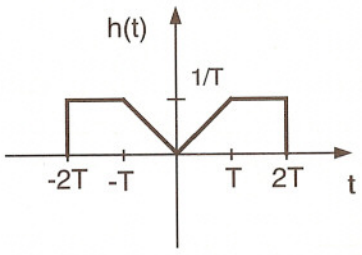


Figura 8.19 Risposta impulsiva del filtro dell'Esempio 8.10

La funzione valor medio $\eta_X(t)$ del processo aleatorio $X(t)$ è data da:

$$\eta_X(t) = E\{X(t)\} = E\left\{\sum_{k=-\infty}^{+\infty} \delta(t - kT - \Theta)\right\}$$

$$= \int_{-\infty}^{+\infty} f_\Theta(\theta) \sum_{k=-\infty}^{+\infty} \delta(t - kT - \theta)\, d\theta = \frac{1}{T} \int_0^T \sum_{k=-\infty}^{+\infty} \delta(t - kT - \theta)\, d\theta$$

$$= \frac{1}{T} \sum_{k=-\infty}^{+\infty} \int_0^T \delta(t - kT - \theta)\, d\theta \qquad \text{(E8.10.2)}$$

Se si effettua il cambiamento di variabile $u = t - kT - \theta$ si ha:

$$\eta_X(t) = -\frac{1}{T} \sum_{k=-\infty}^{+\infty} \int_{t-kT}^{t-(k+1)T} \delta(u)\, du = \frac{1}{T} \sum_{k=-\infty}^{+\infty} \int_{t-(k+1)T}^{t-kT} \delta(u)\, du \qquad \text{(E8.10.3)}$$

L'intervallo di integrazione $(t-(k+1)T, t-kT)$ ricopre, al variare del parametro intero k, tutto l'asse reale senza sovrapposizioni (si veda anche l'Esempio 8.9), per cui

$$\eta_X(t) = \frac{1}{T} \int_{-\infty}^{+\infty} \delta(u)\, du = \frac{1}{T} \qquad \text{(E8.10.4)}$$

Il valor medio del processo $Y(t)$ è allora:

$$\eta_Y(t) \equiv \eta_Y = \eta_X H(0) = \eta_X \int_{-\infty}^{+\infty} h(t)dt = 3 \qquad \text{(E8.10.5)}$$

❑

8.5 Densità spettrale di potenza di un processo stazionario

8.5.1 Definizione e teorema di Wiener-Khintchine

Nel paragrafo precedente abbiamo affrontato e risolto il problema del filtraggio di un processo aleatorio, almeno per quanto riguarda il calcolo delle statistiche semplificate del primo e del secondo ordine. In particolare, abbiamo usato una descrizione del sistema lineare stazionario e dei processi aleatori in ambito *temporale*. Ogniqualvolta si ha a che fare con questioni di filtraggio però può essere conveniente cercare una descrizione *frequenziale* del problema, e questo ci porta all'*analisi di Fourier dei segnali aleatori.*

Ci limiteremo qui ad alcuni cenni di analisi spettrale dei processi aleatori solo nel caso in cui questi ultimi siano *stazionari in senso lato*. L'analisi spettrale dei processi non stazionari non è concettualmente complicata, ma va al di là degli scopi di questo testo.

La caratterizzazione frequenziale dei processi aleatori stazionari in termini di

spettri di ampiezza e fase è poco usuale. Dal punto di vista concettuale, anche un segnale aleatorio può essere scomposto in una sovrapposizione di oscillazioni armoniche, le cui ampiezze e fasi variano parimenti in maniera *aleatoria* al variare della frequenza. Tuttavia, è più comune è limitarsi alla descrizione dello spettro di *potenza* di un processo aleatorio, sul quale ci concentreremo.

Cominciamo con l'osservare che le funzioni campione di un processo stazionario *non possono essere segnali a energia finita*. I segnali a energia finita, infatti, tendono necessariamente a zero quando $t \to \infty$. Se *tutte* le funzioni campione del processo tendessero a zero, necessariamente tenderebbe a zero anche la *funzione valor medio* del processo, che quindi non potrebbe risultare in generale costante (eccetto che per processi a media nulla). Le funzioni campione di un processo stazionario sono segnali in generale a potenza finita, e perciò il segnale aleatorio stesso ammetterà densità spettrale di *potenza*.

Ciò premesso, la definizione di *funzione densità spettrale di potenza* $S_X(f)$ (brevemente, *spettro di potenza*) per un segnale aleatorio è molto simile a quella relativa a un segnale *determinato a potenza finita* $x(t)$ che richiamiamo di seguito:

$$S_x(f) \triangleq \lim_{T\to\infty} \frac{|X_T(f)|^2}{T} \tag{8.5.1}$$

Ricordiamo che $X_T(f)$ è la trasformata di Fourier del segnale $x(t)$ *troncato* sull'intervallo $[-T/2, T/2]$ e quindi ridotto a energia finita. Per un processo aleatorio, si deve pensare di eseguire le stesse operazioni *su ciascuna funzione campione*, ottenendo quindi una quantità *aleatoria* variabile da funzione campione a funzione campione, ovvero con il risultato dell'esperimento:

$$\tilde{S}_x(\omega; f) \triangleq \lim_{T\to\infty} \frac{|\tilde{X}_T(\omega; f)|^2}{T} \tag{8.5.2}$$

ove $\tilde{X}_T(\omega; f)$ è la trasformata di Fourier della generica funzione campione troncata: $\tilde{X}_T(\omega; f) \triangleq \mathcal{F}[x(\omega; t)\text{rect}(t/T)]$. Per ottenere la densità spettrale del *processo*, indipendente cioè dalla particolare funzione campione, bisogna aggiungere a questa definizione un'operazione di *valor medio* per compendiare l'andamento delle particolari $\tilde{S}_x(\omega; f)$ delle varie funzioni campione:

$$S_x(f) \triangleq E\{\tilde{S}_x(\omega; f)\} = \lim_{T\to\infty} \frac{E\{|\tilde{X}_T(f)|^2\}}{T} \tag{8.5.3}$$

Questa definizione è una diretta estensione di quella per segnali determinati, ed è utilizzabile anche per processi *non stazionari*. Sfortunatamente, essa è quasi sempre di difficile applicazione pratica (e comporta anche qualche problema di carattere matematico relativamente all'inversione delle operazioni di valor medio e di limite implicita nella (8.5.3)). Per i processi *stazionari*, si usa allora una *diversa* definizione di densità spettrale di potenza. Ricordiamo infatti il *teorema di Wiener-Khintchine* (si veda il Paragrafo 4.4.3), secondo il quale la densità spettrale di potenza per segnali determinati è comunque calcolabile come trasformata di Fourier della funzione di autocorrelazione $R_X(\tau)$. Questo teorema (la cui dimostrazione è piuttosto lunga e viene omessa) vale anche nel caso di *processi aleatori stazionari*. Possiamo allora *definire* la funzione densità spettrale di potenza $S_X(f)$ di un processo aleatorio $X(t)$ come la trasformata di Fourier della sua funzione di autocorrelazione $R_X(\tau) = E\{X(t)X(t-\tau)\}$:

$$S_X(f) \triangleq \int_{-\infty}^{+\infty} R_X(\tau) e^{-j2\pi f\tau} d\tau = 2\int_{0}^{+\infty} R_X(\tau)\cos(2\pi f\tau) d\tau \tag{8.5.4}$$

ignorando per brevità l'effettiva (ma più complicata) definizione (8.5.3).

Comunque venga definita, la funzione $S_X(f)$ gode delle stesse proprietà elencate a suo tempo per i segnali determinati, e che qui ricapitoliamo:

- $S_X(f)$ è una funzione *reale e pari*: ovvio dalla (8.5.3), ma anche dalla (8.5.4), visto che $S_X(f)$ è la trasformata di Fourier della funzione $R_X(\tau)$, anch'essa reale e pari.
- La potenza media statistica P_X del processo $X(t)$ può essere calcolata integrando $S_X(f)$ su tutto l'asse delle frequenze:

$$P_X = E\{X^2(t)\} = \int_{-\infty}^{+\infty} S_X(f)\, df = 2\int_{0}^{+\infty} S_X(f)\, df \tag{8.5.5}$$

  Questo è ancora ovvio dalla (8.5.3), ma anche dalla definizione alternativa: come per ogni coppia trasformata-antitrasformata, l'integrale in frequenza della $S_X(f)$ è pari al valore per $\tau = 0$ della propria antitrasformata, che è appunto il valore della potenza media statistica: $R_X(0) = E\{X^2(t)\} = P_X$.
- $S_X(f)$ è una funzione *non negativa*:

$$S_X(f) \geq 0 \tag{8.5.6}$$

  Questa proprietà discende immediatamente dalla definizione diretta (8.5.3), mentre è di dimostrazione piuttosto complessa (*teorema di Böchner*) nel caso

di definizione secondo il teorema di Wiener-Khintchine. Dimostreremo questa proprietà nel prossimo paragrafo riconsiderando in ambito frequenziale il *filtraggio* di un processo stazionario.

8.5.2 Filtraggio di un processo aleatorio e densità spettrale di potenza

Riprendiamo dunque il sistema lineare stazionario di Figura 8.18 e cerchiamo di mettere in relazione le caratteristiche *spettrali* dei processi di ingresso $X(t)$ e di uscita $Y(t)$, entrambi *stazionari in senso lato*. La densità spettrale di potenza $S_Y(f)$ di quest'ultimo è

$$S_Y(f) = \mathcal{F}\left[R_Y(\tau)\right] = \mathcal{F}\left[R_X(\tau) \otimes h(\tau) \otimes h(-\tau)\right] = S_X(f)H(f)H(-f) \qquad (8.5.7)$$

Poiché la risposta impulsiva $h(t)$ del sistema è un segnale reale, la sua trasformata gode della proprietà di simmetria Hermitiana ($H^*(f) = H(-f)$), e la (8.5.7) diventa

$$S_Y(f) = S_X(f)H(f)H^*(f) = S_X(f)|H(f)|^2 \qquad (8.5.8)$$

cioè esattamente la relazione (4.4.22) che vale anche per segnali *determinati*. Lo spettro di potenza del processo di uscita $Y(t)$ può essere ricavato da quello del processo d'ingresso note le caratteristiche di selettività in frequenza del sistema, che sono riassunte nella *risposta in ampiezza* al quadrato $|H(f)|^2$. Ancora una volta, la risposta in *fase* nel sistema non influenza il contenuto di potenza del processo di uscita. Osserviamo incidentalmente che la potenza media statistica P_Y del processo di uscita $Y(t)$ può essere calcolata in ambito frequenziale come segue:

$$P_Y = R_Y(0) = \int_{-\infty}^{+\infty} S_Y(f)df = \int_{-\infty}^{+\infty} S_X(f)|H(f)|^2 df = 2\int_{0}^{+\infty} S_X(f)|H(f)|^2 df \qquad (8.5.9)$$

La relazione del filtraggio (8.5.8) è importante perché permette di dimostrare che i) la densità spettrale di potenza di un processo stazionario è una funzione non negativa e ii) la funzione $S_X(f)$, definita come trasformata di Fourier della $R_X(\tau)$, descrive la distribuzione della potenza sulle varie componenti frequenziali nello spettro del segnale aleatorio $X(t)$. Riprendendo il ragionamento fatto a questo proposito nel Paragrafo 4.4, filtriamo $X(t)$ con il filtro passa-banda ideale di Figura 8.20. La funzione $S_X(f)$ del segnale di ingresso è quella rappresentata a tratto sottile in Figura 8.21, mentre l'andamento della corrispondente $S_Y(f)$ è disegnato a tratto spesso. La potenza del processo di uscita $Y(t)$ si può calcolare come segue:

$$P_Y = \int_{-\infty}^{\infty} S_Y(f)\, df = \int_{-\infty}^{\infty} S_X(f)\ |H(f)|^2\, df = 2 \int_{-\bar{f}-\Delta f/2}^{-\bar{f}+\Delta f/2} S_X(f)\ df \tag{8.5.10}$$

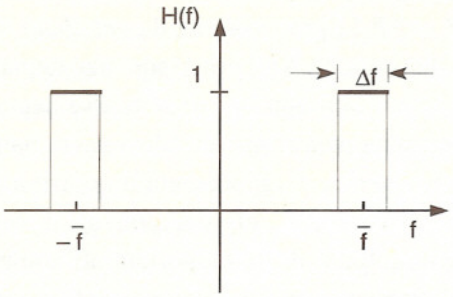


Figura 8.20 Filtro passa-banda ideale con frequenza centrale $\bar{f}$ e banda Δf

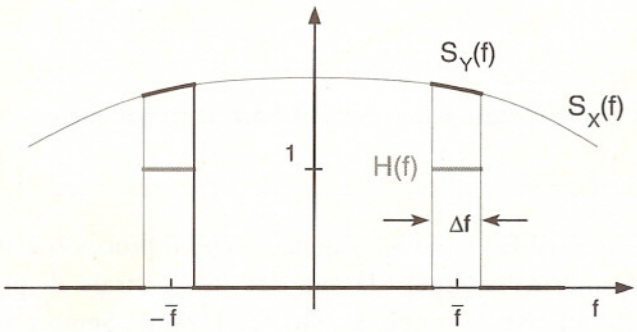


Figura 8.21 Andamento delle funzioni $S_x(f)$ ed $S_y(f)$

Se si riduce progressivamente la banda passante Δf del filtro, cioè si considera un filtro estremamente *selettivo*, si può approssimare $S_X(f)$ all'interno della banda stessa con una *costante* di valore pari a $S_X(\bar{f})$. Allora la potenza del segnale d'uscita può essere approssimata da

$$P_Y \cong 2\Delta f \cdot S_X(\bar{f}) \tag{8.5.11}$$

e quindi

$$S_X(\bar{f}) \cong \frac{1}{2}\frac{P_Y}{\Delta f} \triangleq \frac{1}{2}\frac{\Delta P_X(\bar{f})}{\Delta f} \tag{8.5.12}$$

Stante l'arbitrarietà di $\bar{f}$, la funzione $S_X(\cdot)$ deve essere positiva o al limite nulla per tutti i valori della frequenza, in quanto P_Y è non negativa. Notiamo, inoltre, che $Y(t)$ è stato ottenuto da $X(t)$ sopprimendo tutte le componenti frequenziali al di fuori di un intorno della frequenza $\pm\bar{f}$. Dunque, come suggerisce la notazione nella (4.4.12), $P_Y = \Delta P_X(\bar{f})$ rappresenta il *contributo* alla potenza totale P_X delle sole componenti del segnale $X(t)$ con frequenze appartenenti all'intervallo $[\bar{f} - \Delta f/2, \bar{f} + \Delta f/2\,]$ (più il simmetrico sul semiasse negativo), in quanto tutte le altre componenti sono state *cancellate* dal filtro passa-banda. Allora il particolare valore $S_X(\bar{f})$ alla frequenza $\bar{f}$ rappresenta il contributo *locale* alla potenza totale del segnale $X(t)$ dovuto alle *sole* componenti di segnale con frequenza appartenente a un piccolo intorno di $\bar{f}$, rapportato all'ampiezza piccola Δf dell'intervallo stesso. Ciò corrisponde alla classica definizione di *densità* di una grandezza fisica (la potenza) rispetto a una misura di estensione (qui l'ampiezza di una banda), e giustifica il nome con il quale si designa la funzione $S_X(f)$: *densità spettrale di potenza* del segnale $X(t)$.

Esempio 8.11

Calcoliamo la densità spettrale di potenza del processo parametrico

$$X(t) = a\cos(2\pi f_0 t + \Theta) \tag{E8.11.1}$$

ove $\Theta \in \mathcal{U}(0, 2\pi)$. Dall'Esempio 8.5 sappiamo che il processo è stazionario in senso lato, e quindi calcoleremo la sua densità spettrale di potenza come trasformata della relativa funzione di autocorrelazione. Sempre nell'Esempio 8.5, abbiamo trovato che

$$R_X(\tau) = \frac{a^2}{2}\cos(2\pi f_0 \tau) \tag{E8.11.2}$$

per cui la densità spettrale di potenza $S_X(f)$ di $X(t)$ è (Figura 8.22):

$$S_X(f) = \mathcal{F}[R_X(\tau)] = \mathcal{F}\left[\frac{a^2}{2}\cos(2\pi f_0 \tau)\right] = \frac{a^2}{4}\left[\delta(f - f_0) + \delta(f + f_0)\right] \tag{E8.11.3}$$

La potenza del processo è quindi *concentrata* sulle due frequenze $\pm f_0$. ❑

8.5.3 Processo di rumore bianco

Nel Paragrafo 8.3 abbiamo introdotto la nozione di *tempo di correlazione* per misurare la rapidità media di variazione delle funzioni campione di un processo.

La corrispondente grandezza in ambito frequenziale è naturalmente la *banda* dello spettro di potenza del processo, che dà la stessa indicazione del tempo di correlazione. Se la funzione di autocorrelazione di un processo decresce rapidamente, cioè il tempo di correlazione è piccolo, la densità spettrale corrispondente ha una banda grande, come illustrato nelle Figure 8.23a-8.23b, viceversa se il tempo di correlazione è grande. Quindi, come era lecito aspettarsi, quanto maggiore è la rapidità di variazione delle realizzazioni di un processo, tanto più grande è la banda del suo spettro di potenza. La Figura 8.24 mostra tre funzioni campione di tre processi aleatori $X_1(t)$, $X_2(t)$ e $X_3(t)$ con banda progressivamente crescente; si nota che la rapidità di variazione cresce, così come l'ampiezza delle escursioni del segnale (per effetto dell'incremento della potenza del segnale stesso).

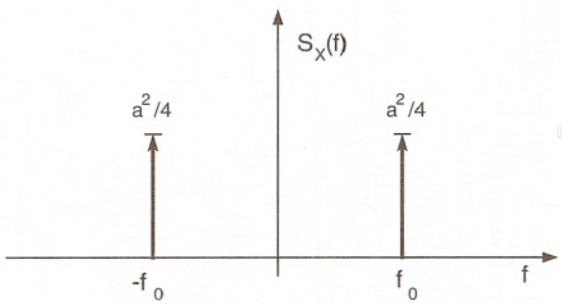


Figura 8.22 Spettro di potenza del processo nell'Esempio 8.11

Cosa succede se consideriamo una situazione come quella di Figura 8.23b, in cui la banda dello spettro di potenza tende a crescere *illimitatamente*, mantenendo lo spettro sempre il medesimo valore per $f = 0$? Evidentemente, la densità spettrale di potenza del processo $X(t)$ tende a diventare *costante* (di valore, diciamo, $S_X(0) = \xi$) mentre il tempo di correlazione τ_{cor} tende a ridursi sempre più. Al limite, si arriva a una situazione in cui la funzione di autocorrelazione è *impulsiva* (Figura 8.25):

$$S_X(f) = S_X(0) = \xi \Leftrightarrow R_X(\tau) = \xi\delta(\tau) \tag{8.5.13}$$

Un processo aleatorio stazionario (almeno) in senso lato che presenta queste caratteristiche statistiche viene chiamato *δ-correlato* o più comunemente, nell'ingegneria, *processo di rumore bianco*. L'appellativo *bianco* deriva dall'analogia dello spettro di potenza di questo processo con quello della *luce bianca*: il ru-

more bianco contiene componenti a *tutte* le frequenze, da $-\infty$ a $+\infty$, con la stessa intensità, così come la luce bianca "contiene tutti i colori" (si ricordi l'esperimento del *disco di Newton*). Osserviamo inoltre che un processo bianco ha sempre valor medio nullo essendo, per la (8.3.19), $\lim_{\tau\to\infty} R_N(\tau) = \eta_N^2 = 0$.

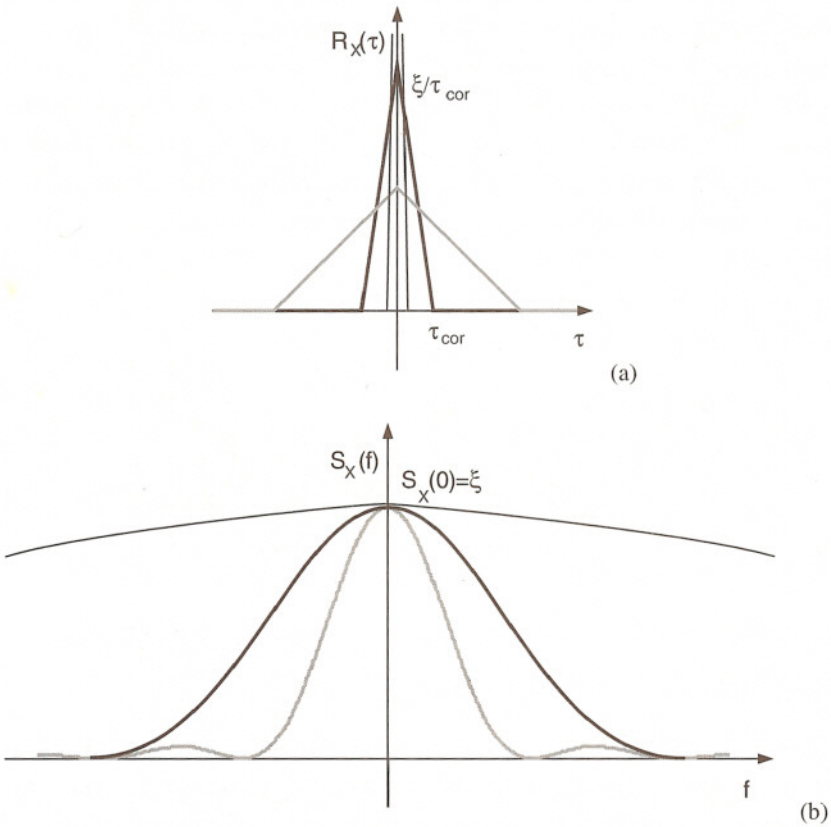


Figura 8.23 Funzione di autocorrelazione e densità spettrale di potenza

Evidentemente, un processo bianco è solo un'*astrazione matematica*: lo spettro di potenza *costante* (8.5.13) comporta che la potenza di questo segnale sia *infinita*, condizione impossibile per un segnale fisico. È anche problematico *rappresentare* la funzione campione di un processo bianco, che deve intendersi come "caso-limite" della funzione campione in Figura 8.23c, pensando di aumentarne ulteriormente (e illimitatamente) l'ampiezza e la velocità di variazione. Nonostante ciò, il rumore bianco è usato molto frequentemente nell'ingegneria come *modello* per una varietà di segnali coinvolti in molti fenomeni fisici, come testimonia l'esempio seguente.

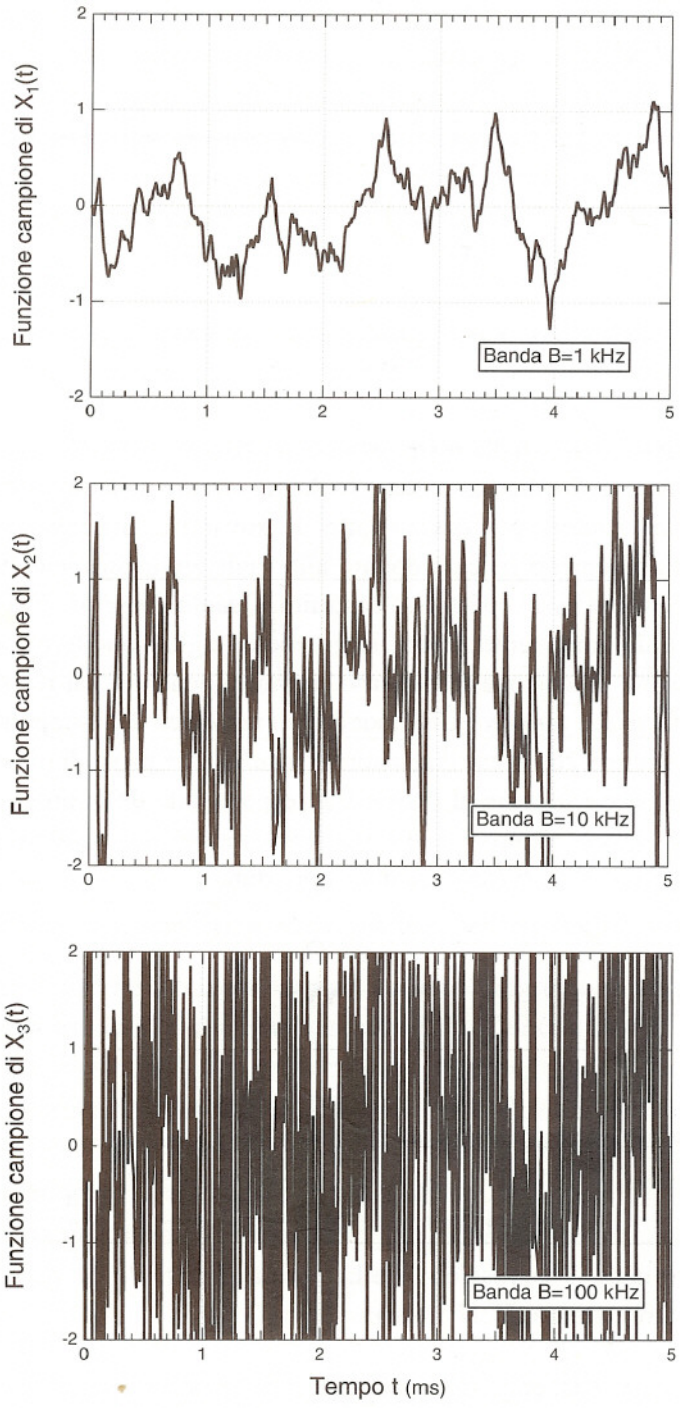


Figura 8.24 Funzioni campione di processi aventi diverse bande dello spettro di potenza

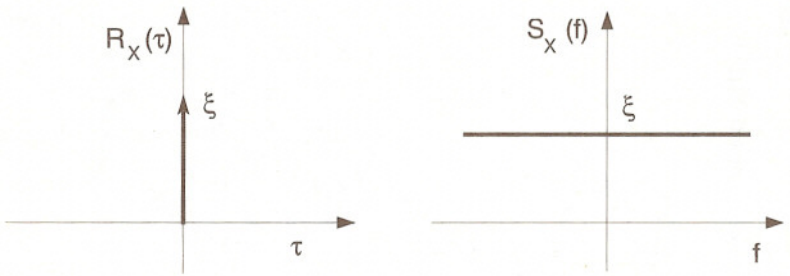


Figura 8.25 Funzione di autocorrelazione e densità spettrale di potenza di un processo bianco

Esempio 8.12

Un comune resistore usato nella realizzazione dei circuiti elettronici, oltre a presentare una resistenza R (ohm) al passaggio della corrente, genera anche una debole *tensione di rumore* per il solo fatto di trovarsi a una *temperatura* T_R (kelvin). Questo fenomeno fisico è dovuto alla "agitazione termica" degli elettroni del materiale con cui il resistore è costruito. Quanto maggiore è la temperatura cui si trova il componente, tanto più grande è l'agitazione termica, e tanto più grande è anche la tensione di disturbo (rumore) generata dal resistore, che viene chiamata *rumore termico*. Un modello più realistico del componente è allora quello di Figura 8.26, in cui il resistore è *ideale* (cioè privo di disturbo), e il generatore di tensione in serie al resistore è responsabile della produzione del rumore termico. Quest'ultimo viene a sua volta modellato come un processo aleatorio stazionario $N(t)$ di caratteristiche opportune.

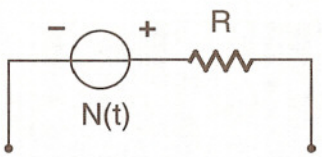


Figura 8.26 Resistore con generatore di rumore termico

La descrizione del fenomeno del rumore termico è abbastanza complessa, e coinvolge considerazioni di meccanica quantistica. Attraverso quest'analisi, si arriva a determinare l'espressione della *densità spettrale di potenza* della tensione di rumore termico (H. Nyquist):

$$S_N(f) = 2kT_R R \frac{|f|/f_0}{\exp(|f|/f_0) - 1} \ (\mathrm{V^2/Hz}) \quad , \quad f_0 = \frac{kT_R}{h} \ (\mathrm{Hz}) \qquad \text{(E8.12.1)}$$

ove k è la costante di Boltzmann ($k = 1.38 \cdot 10^{-23}$ J/K) e h è la costante di Planck ($h = 6.62 \cdot 10^{-34}$ J·s). Alla temperatura ambiente ($T_R = 290$ K) la frequenza f_0 caratteristica dello spettro (E8.12.1) è pari circa a $f_0 = 6.05$ THz, cioè $f_0 = 6050$ GHz! L'andamento normalizzato di questa densità spettrale è mostrato in Figura 8.27; si vede che per frequenze molto più piccole di f_0 lo spettro di potenza del rumore termico è praticamente *piatto* (costante) e vale

$$S_N(f) \cong S_N(0) = 2kT_R R \qquad \text{(E8.12.2)}$$

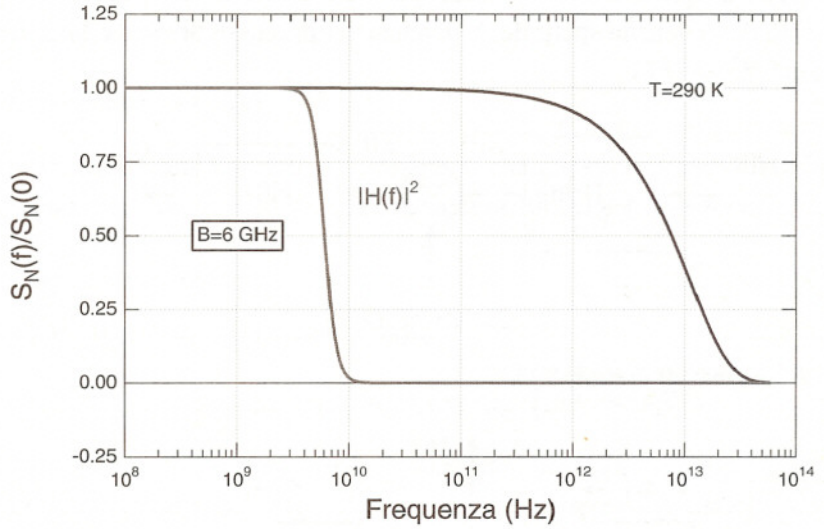


Figura 8.27 Densità spettrale del rumore termico piatta nella banda di un filtro

Supponiamo ora che il rumore termico si trovi all'ingresso di un qualche sistema filtrante con banda B. Nella grande maggioranza dei casi pratici, la banda B del filtro sarà *di alcuni ordini di grandezza* più piccola di f_0. Questo spiega perché il rumore bianco è un modello molto utile: come è suggerito dalla Figura 8.27 stessa, per calcolare gli effetti del rumore termico sull'*uscita* di un qualunque sistema filtrante, l'effettiva densità di potenza (E8.12.1) del rumore termico $N(t)$ può tranquillamente essere sostituita da quella di un (fittizio) rumore bianco $W(t)$:

$$S_W(f) = S_N(0) = 2kT_R R = \xi \qquad \text{(E8.12.3)}$$

In entrambi i casi (reale e fittizio) il sistema "vede" la stessa densità piatta nella sua propria banda, ma, naturalmente, ogni tipo di calcolo è semplificato con

questo approccio, perché la densità di potenza del rumore bianco $S_W(f) = \xi$ è assai più semplice di quella effettiva del rumore termico $S_N(f)$. ❑

Esempio 8.13

Nel sistema illustrato nella Figura 8.28, il processo aleatorio $N(t)$ è stazionario in senso lato con densità spettrale di potenza $S_N(f) = \xi$ (rumore bianco), Θ è una variabile aleatoria indipendente da $N(t)$ e uniformemente distribuita nell'intervallo $[-\pi, \pi)$, e infine il filtro è un passa-banda ideale la cui risposta in frequenza $H(f)$ è mostrata nella Figura 8.29. Determiniamo e rappresentiamo l'andamento della densità spettrale di potenza del processo di uscita $Y(t)$, nell'ipotesi in cui $f_0 >> 1/T$.

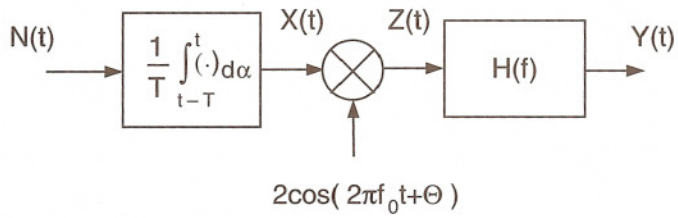


Figura 8.28 Sistema dell'Esempio 8.13

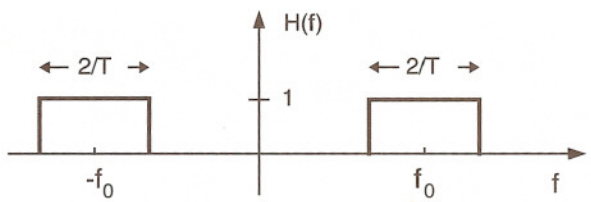


Figura 8.29 Risposta in frequenza del filtro indicato nello schema della Figura 8.28

Cominciamo con l'osservare che il processo $X(t)$ è stazionario in senso lato, perché risulta dal filtraggio con un SLS di un processo bianco, quindi stazionario in senso lato. Il sistema integratore a finestra mobile ha risposta impulsiva

$$h_0(t) = \frac{1}{T}\text{rect}\left(\frac{t - T/2}{T}\right) \qquad \text{(E8.13.1)}$$

per cui la funzione valor medio $\eta_X(t)$ di $X(t)$ è

$$\eta_X = \eta_N H(0) = 0 \qquad \text{(E8.13.2)}$$

in quanto la funzione valor medio $\eta_N(t)$ di $N(t)$ è identicamente nulla (processo

bianco). La funzione di autocorrelazione $R_X(\tau)$ di $X(t)$ è data, invece, dalla seguente relazione:

$$R_X(\tau) = R_N(\tau) \otimes h_0(\tau) \otimes h_0(-\tau) = \xi\delta(\tau) \otimes R_{h_0}(\tau) = \frac{\xi}{T}\left(1 - \frac{|\tau|}{T}\right)\text{rect}\left(\frac{\tau}{2T}\right) \tag{E8.13.3}$$

La corrispondente densità spettrale di potenza è

$$S_X(f) = \mathcal{F}\left[R_X(\tau)\right] = \xi\,\text{sinc}^2(fT) \tag{E8.13.4}$$

Chiamiamo ora $S(t)$ il processo

$$S(t) = 2\cos(2\pi f_0 t + \Theta) \tag{E8.13.5}$$

in Figura 8.18. Dall'Esempio 8.5 sappiamo che questo processo è stazionario in senso lato con valor medio η_S *nullo* e funzione di autocorrelazione

$$R_S(\tau) = 2\cos(2\pi f_0 \tau) \tag{E8.13.6}$$

Calcoliamo ora la funzione valor medio di $Z(t)$:

$$\eta_Z(t) = E\{Z(t)\} = E\{N(t)S(t)\} = 2E\{X(t)\cos(2\pi f_0 t + \Theta)\} \tag{E8.13.7}$$

Notiamo che la variabile aleatoria Θ, essendo indipendente dal processo $N(t)$, è anche indipendente da $X(t)$. Quindi, fissato il tempo t, anche la variabile aleatoria $\cos(2\pi f_0 t + \Theta)$ è indipendente da $X(t)$. In conclusione:

$$\eta_Z(t) = E\{X(t)\}E\{S(t)\} = 0 \tag{E8.13.8}$$

La funzione di autocorrelazione $R_Z(t, t-\tau)$ è data da

$$\begin{aligned} R_Z(t,t-\tau) &= E\{Z(t)Z(t-\tau)\} \\ &= 4E\{X(t)X(t-\tau)\cos(2\pi f_0 t + \Theta)\cos(2\pi f_0(t-\tau) + \Theta)\} \\ &= 4E\{X(t)X(t+\tau)\}E\{\cos(2\pi f_0 t + \Theta)\cos(2\pi f_0(t-\tau) + \Theta)\} \\ &= 2R_X(\tau)\cos(2\pi f_0 \tau) = R_Z(\tau) \end{aligned} \tag{E8.13.9}$$

dove abbiamo ancora una volta sfruttato l'indipendenza da Θ di *tutte* le variabili aleatorie estratte dal processo $X(t)$.

Il risultato principale di questo calcolo è che il processo $Z(t)$ è stazionario in senso lato, avendo valor medio nullo e funzione di autocorrelazione (E8.13.9)

che *non* dipende da t. La densità spettrale di potenza di $Z(t)$ è

$$S_Z(f) = \mathcal{F}[R_Z(\tau)] = S_X(f - f_0) + S_X(f + f_0)$$
$$= \xi \left\{ \text{sinc}^2[(f + f_0)T] + \text{sinc}^2[(f - f_0)T] \right\} \qquad \text{(E8.13.10)}$$

Se vale la condizione $f_0 >> 1/T$, le due repliche della $S_X(f)$ che compaiono nella $S_Z(f)$ centrate a $\pm f_0$ si possono considerare in pratica "isolate", cioè non interferenti reciprocamente: le "code" di ciascuna delle due funzioni $\text{sinc}^2(\cdot)$ si smorzano completamente prima di raggiungere l'altra, come (approssimativamente) in Figura 8.30.

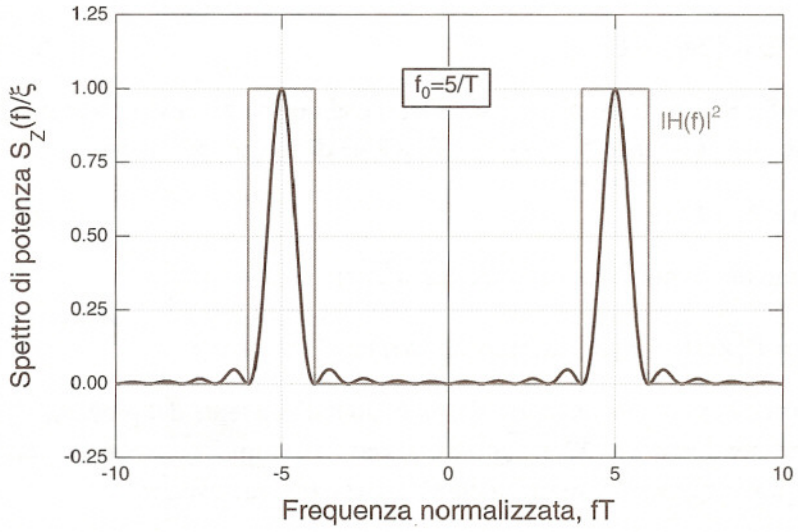


Figura 8.30 Densità spettrale di potenza $S_Z(f)$ dell'Esempio 8.13

La densità $S_Y(f)$ del processo $Y(t)$ si ricava infine dal teorema fondamentale del filtraggio (8.5.8):

$$S_Y(f) = S_Z(f)|H(f)|^2 \qquad \text{(E8.13.11)}$$

e la condizione $f_0 >> 1/T$ permette di concludere che

$$S_Y(f) \cong \xi \left[\text{sinc}^2((f - f_0)T)\,\text{rect}\left(\frac{f - f_0}{2/T}\right) + \text{sinc}^2((f + f_0)T)\,\text{rect}\left(\frac{f + f_0}{2/T}\right) \right] \qquad \text{(E8.13.12)}$$

L'andamento della densità spettrale di potenza $S_Y(f)$ è illustrato nella Figura 8.31. Si nota che il filtro $H(f)$ ha selezionato le sole componenti frequenziali relative ai *lobi principali* delle funzioni $\text{sinc}^2(\cdot)$ nello spettro di potenza di $Z(t)$.

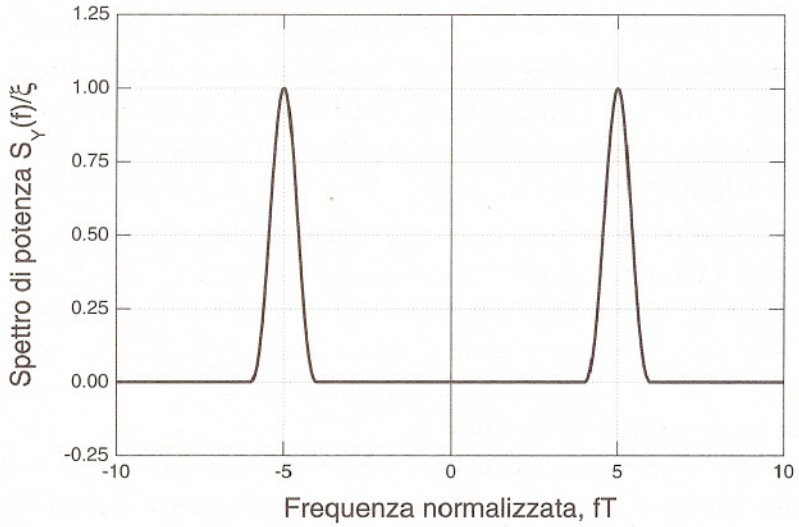


Figura 8.31 Densità spettrale di potenza $S_Y(f)$ dell'Esempio 8.13

❑

8.6 Processi aleatori Gaussiani

8.6.1 Definizione e prime proprietà

Nell'Esempio 8.12 abbiamo esaminato alcune proprietà del *rumore termico*. In particolare, ne abbiamo studiato la densità spettrale di potenza e la funzione di autocorrelazione, cioè le caratteristiche *temporali* e *spettrali*, senza specificarne le caratteristiche di *distribuzione delle ampiezze*. Fissiamo dunque un istante t_1 ed estraiamo dal processo la variabile aleatoria $X(t_1)$. Possiamo osservare che il particolare valore del rumore termico ottenuto in una prova dell'esperimento a quel dato istante è determinato da un gran numero di contributi elementari *indipendenti* di tensione di rumore, provocati ciascuno dai singoli elettroni in agitazione termica all'interno del resistore. Questa osservazione giustifica l'applicazione del teorema-limite centrale 7.4.9 e permette di concludere che le statistiche di $X(t_1)$ sono con ottima approssimazione *Gaussiane*. Ma possiamo dire anche di più: con considerazioni simili, si trova che la n-upla di variabili $[X(t_1), X(t_2), ..., X(t_n)]$ estratte agli istanti $(t_1, t_2, ..., t_n)$ risulta un sistema di

variabili aleatorie *congiuntamente Gaussiane*.

Questa proprietà, verificata "sperimentalmente" da un particolare segnale generato da un altrettanto particolare fenomeno fisico, può essere generalizzata in modo da definire un'intera *classe* di processi aleatori che modellano una quantità di fenomeni fisici anche diversi dal rumore termico: i *processi Gaussiani*.

Definizione Un processo aleatorio $X(t)$ è *Gaussiano* se le n variabili aleatorie $[X(t_1), X(t_2),..., X(t_n)]$ da esso estratte agli istanti $(t_1, t_2,..., t_n)$ risultano *congiuntamente Gaussiane* comunque si scelga il valore del parametro intero n e per qualunque n-upla di istanti $(t_1, t_2,..., t_n)$.

Questa definizione è piuttosto restrittiva: non basta che *una* variabile aleatoria $X(t_1)$ estratta a un certo istante t_1 sia Gaussiana per poter dire che il *processo* è Gaussiano. La definizione dice infatti che un *vettore* aleatorio riempito con un numero *qualunque* (diremmo, arbitrariamente grande) di variabili estratte dal processo a istanti *qualunque* deve comunque essere un vettore aleatorio Gaussiano (si veda il Paragrafo 7.4.8).

Molti segnali incontrati in altrettanti fenomeni fisici possono essere modellati come realizzazioni di un processo Gaussiano (segnali sismici, segnali vocali e musicali, segnali radio ecc.). Questo spiega la "centralità" di questa classe di processi, e la necessità di studiarne le proprietà in maniera approfondita.

Cominciamo con l'osservare che un processo aleatorio Gaussiano $X(t)$ è completamente caratterizzato dal punto di vista statistico quando sono note la sua funzione valor medio $\eta_X(t)$ e la sua funzione di autocorrelazione $R_X(t_1, t_2)$. Ricordiamo che la descrizione statistica di un processo $X(t)$ è *completa* se è nota la *densità di probabilità di ordine n* $f_X(x_1, x_2,..., x_n; t_1, t_2,..., t_n)$ (8.1.9), comunque grande sia il numero intero n e comunque si scelga la n-upla di istanti $(t_1, t_2..., t_n)$. Indichiamo dunque con $X_1,..., X_n$ le n variabili aleatorie estratte dal processo: $X_1 = X(t_1), X_2 = X(t_2),..., X_n = X(t_n)$. Se il processo è Gaussiano, la densità congiunta di queste n variabili è *per definizione* Gaussiana, cioè ha la forma (7.4.49):

$$f_X(x_1, x_2,..., x_n; t_1, t_2,..., t_n) = \frac{1}{\sqrt{(2\pi)^n \det \mathbf{C}_\mathbf{X}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\eta}_\mathbf{X})^T \mathbf{C}_\mathbf{X}^{-1}(\mathbf{x} - \boldsymbol{\eta}_\mathbf{X})\right] \tag{8.6.1}$$

ove $\mathbf{x} = (x_1, x_2,..., x_n)^T$. Le grandezze $\boldsymbol{\eta}_\mathbf{X}$ e $\mathbf{C}_\mathbf{X}$ possono essere immediatamente esplicitate non appena si conosca l'andamento delle funzioni $\eta_X(t)$ e $R_X(t_1, t_2)$. Infatti il vettore dei valori medi $\boldsymbol{\eta}_\mathbf{X}$ è dato da

$$\eta_{\mathbf{X}} \triangleq \left[E\{X_1\}, E\{X_2\}, \ldots, E\{X_n\}\right]^T = \left[E\{X(t_1)\}, E\{X(t_2)\}, \ldots, E\{X(t_n)\}\right]^T$$
$$= \left[\eta_X(t_1), \eta_X(t_2), \ldots, \eta_X(t_n)\right]^T \tag{8.6.2}$$

e quindi (com'è ovvio) i suoi elementi sono *i valori della funzione valor medio del processo* agli istanti $(t_1, t_2 \ldots, t_n)$ di estrazione delle variabili aleatorie. Per quanto riguarda la matrice di covarianza $\mathbf{C_X}$, il suo generico elemento c_{ik} è dato da

$$c_{ik} \triangleq E\left\{\left[X_i - \eta_{X_i}\right]\left[X_k - \eta_{X_k}\right]\right\} = E\left\{\left[X(t_i) - \eta_X(t_i)\right]\left[X(t_k) - \eta_X(t_k)\right]\right\}$$
$$= C_X(t_i, t_k) = R_X(t_i, t_k) - \eta_X(t_i)\eta_X(t_k) \tag{8.6.3}$$

e quindi

$$\mathbf{C_X} = \begin{bmatrix} C_X(t_1,t_1) & C_X(t_1,t_2) & \ldots & C_X(t_1,t_n) \\ C_X(t_2,t_1) & C_X(t_2,t_2) & \ldots & C_X(t_2,t_n) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(t_n,t_1) & C_X(t_n,t_2) & \ldots & C_X(t_n,t_n) \end{bmatrix} \tag{8.6.4}$$

ove $C_X(\cdot,\cdot)$ è la *funzione di autocovarianza* di $X(t)$. La matrice di covarianza è quindi nota non appena è nota la funzione di autocovarianza del processo o, equivalentemente, la funzione di autocorrelazione e la funzione valor medio. Dal vettore dei valori medi e dalla matrice di covarianza si ricava poi la densità di probabilità del processo come nella (8.6.1) per qualunque valore di n e di $(t_1, t_2 \ldots, t_n)$: la caratterizzazione del processo è *completa*.

La particolare forma della densità Gaussiana comporta un'ulteriore proprietà dei processi Gaussiani: se un processo Gaussiano $X(t)$ è stazionario in senso lato, allora è anche *stazionario in senso stretto*. Se infatti il processo è stazionario in senso lato, la funzione valor medio è costante ($\eta_X(t) \equiv \eta_X$) e la funzione di autocovarianza dipende solo dalla differenza dei due istanti considerati ($C_X(t_1, t_2) = C_X(t_1 - t_2)$). In tal caso, il vettore dei valori medi $\boldsymbol{\eta}_{\mathbf{X}}$ (8.6.2) è

$$\boldsymbol{\eta}_{\mathbf{X}} = \left[\eta_X, \eta_X, \ldots, \eta_X\right]^T = \eta_X \mathbf{1} \tag{8.6.5}$$

e la matrice di covarianza $\mathbf{C_X}$ diventa

$$\mathbf{C}_X = \begin{bmatrix} C_X(0) & C_X(t_1 - t_2) & \dots & C_X(t_1 - t_n) \\ C_X(t_2 - t_1) & C_X(0) & \dots & C_X(t_2 - t_n) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(t_n - t_1) & C_X(t_n - t_2) & \dots & C_X(0) \end{bmatrix} \tag{8.6.6}$$

Consideriamo, adesso, la n-upla di istanti $(t_1 + \Delta t, t_2 + \Delta t, ..., t_n + \Delta t)$ ottenuta per *traslazione rigida* dalla n-upla originaria, e ricalcoliamo la densità congiunta del processo per questi nuovi istanti. Ciò equivale a ricalcolare il vettore dei valori medi e la matrice di covarianza, perché la densità dipende soltanto da queste grandezze. Il vettore $\boldsymbol{\eta}_X$ (8.6.5) *non dipende dal tempo* e resta quindi invariato. La nuova matrice di covarianza è

$$\mathbf{C}'_X = \begin{bmatrix} C_X(0) & C_X(t_1 + \Delta t - (t_2 + \Delta t)) & \dots & C_X(t_1 + \Delta t - (t_n + \Delta t)) \\ C_X(t_2 + \Delta t - (t_1 + \Delta t)) & C_X(0) & \dots & C_X(t_2 + \Delta t - (t_n + \Delta t)) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(t_n + \Delta t - (t_1 + \Delta t)) & C_X(t_n + \Delta t - (t_2 + \Delta t)) & \dots & C_X(0) \end{bmatrix}$$

$$= \begin{bmatrix} C_X(0) & C_X(t_1 - t_2) & \dots & C_X(t_1 - t_n) \\ C_X(t_2 - t_1) & C_X(0) & \dots & C_X(t_2 - t_n) \\ \vdots & \vdots & \ddots & \vdots \\ C_X(t_n - t_1) & C_X(t_n - t_2) & \dots & C_X(0) \end{bmatrix} = \mathbf{C}_X \tag{8.6.7}$$

Poiché il vettore dei valori medi e la matrice di covarianza *non sono cambiate* a causa della traslazione temporale, concludiamo che anche la *funzione densità* è invariata:

$$f_X(x_1, x_2, ..., x_n; t_1 + \Delta t, t_2 + \Delta t, ..., t_n + \Delta t) = f_X(x_1, x_2, ..., x_n; t_1, t_2, ..., t_n) \tag{8.6.8}$$

comunque si scelga la traslazione Δt. Quindi il processo Gaussiano e stazionario in senso lato $X(t)$ è anche *stazionario in senso stretto*.

8.6.2 Filtraggio dei processi Gaussiani

Come abbiamo visto nel Paragrafo 8.4, il problema del *filtraggio* di un processo aleatorio non è completamente risolubile, nel senso che è in generale impossibile ottenere la descrizione completa del processo d'uscita $Y(t)$ nota quella del processo d'ingresso $X(t)$. Supponiamo però che $X(t)$ sia Gaussiano: all'uscita del filtro (SLS) avente risposta impulsiva $h(t)$ avremo il processo

$$Y(t) = X(t) \otimes h(t) = \int_{-\infty}^{+\infty} X(\alpha) h(t-\alpha) d\alpha \tag{8.6.9}$$

ove la convoluzione è da intendersi come precisato nel Paragrafo 8.4. Questo integrale di convoluzione può essere *approssimato* con una sommatoria:

$$Y(t) = \int_{-\infty}^{+\infty} X(\alpha) h(t-\alpha) d\alpha \cong \sum_{k=-\infty}^{+\infty} X(k\Delta\alpha) h(t-k\Delta\alpha)\Delta\alpha \tag{8.6.10}$$

e l'approssimazione è tanto più accurata quanto più piccolo è il periodo di campionamento $\Delta\alpha$ del nostro segnale. Fissiamo adesso la consueta n-upla di istanti $(t_1, t_2, ..., t_n)$ e consideriamo le n variabile aleatorie estratte dal processo di uscita a questi istanti:

$$\begin{aligned} Y(t_1) &\cong \sum_{k=-\infty}^{+\infty} X(k\Delta\alpha) h(t_1 - k\Delta\alpha)\Delta\alpha \\ Y(t_2) &\cong \sum_{k=-\infty}^{+\infty} X(k\Delta\alpha) h(t_2 - k\Delta\alpha)\Delta\alpha \\ &\vdots \\ Y(t_n) &\cong \sum_{k=-\infty}^{+\infty} X(k\Delta\alpha) h(t_n - k\Delta\alpha)\Delta\alpha \end{aligned} \tag{8.6.11}$$

Nota la risposta impulsiva del sistema, le varie quantità $h(t_i - k\Delta\alpha)\Delta\alpha$ con i e k variabili sono dei *coefficienti determinati*: $a_{ik} = h(t_i - k\Delta\alpha)\Delta\alpha$, e quindi la (8.6.11) diventa:

$$\begin{aligned} Y(t_1) &\cong \sum_{k=-\infty}^{+\infty} a_{1k} X(k\Delta\alpha) \\ Y(t_2) &\cong \sum_{k=-\infty}^{+\infty} a_{2k} X(k\Delta\alpha) \\ &\vdots \\ Y(t_n) &\cong \sum_{k=-\infty}^{+\infty} a_{nk} X(k\Delta\alpha) \end{aligned} \quad \Rightarrow \quad \mathbf{Y} = \mathbf{AX} \tag{8.6.12}$$

dove $\mathbf{A} = \{a_{ik}\}$, $\mathbf{Y} = \left[Y(t_1), Y(t_2), \ldots, Y(t_n)\right]^T$ è il vettore delle variabili estratte dal processo d'uscita e $\mathbf{X} = \left[\ldots, X((k-1)\Delta\alpha), X(k\Delta\alpha), X((k+1)\Delta\alpha), \ldots\right]^T$ contiene variabili aleatorie estratte dal processo d'ingresso. Dalla definizione di processo Gaussiano, sappiamo che $\mathbf{X}$ è un *vettore Gaussiano*; vediamo inoltre che la (8.6.12) è una *trasformazione lineare* tra vettori aleatori. La conclusione è che *anche* $\mathbf{Y}$ *è un vettore Gaussiano* qualunque sia la scelta di n e degli istanti

$(t_1, t_2, ..., t_n)$: il processo $Y(t)$ è esso stesso Gaussiano.

I processi Gaussiani sono l'eccezione che conferma la regola, nel senso che per questi è possibile dare una descrizione statistica completa del processo all'uscita di un SLS, quando siano note le caratteristiche del processo d'ingresso. La proprietà di "conservazione della Gaussianità" che abbiamo dimostrato per un sistema lineare stazionario è valida anche per *sistemi lineari ma non stazionari.* Come è chiaro, se il processo (Gaussiano) d'ingresso a un SLS è stazionario in senso lato (e quindi anche in senso stretto), allora il processo di uscita è anch'esso stazionario. Se il sistema lineare *non* è stazionario, il processo di uscita è ancora Gaussiano ma in generale perde la proprietà di stazionarietà.

Esempio 8.14 (Processo di Ornstein-Uhlenbeck)
Calcoliamo la funzione di autocorrelazione, lo spettro di potenza e la densità di probabilità del primo ordine $f_Y(y;t)$ del processo $Y(t)$ in uscita alla squadra *R-C* di Figura 8.32 ($\alpha = RC$) quando il segnale d'ingresso è un processo *Gaussiano bianco* $X(t)$ con densità spettrale di potenza $S_X(f) = N_0/2$.

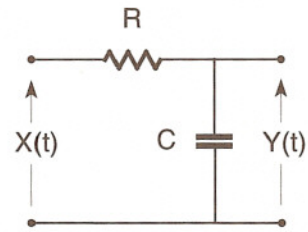


Figura 8.32 Processo Gaussiano bianco in ingresso alla squadra *R-C*

Dalla definizione generale di processo bianco, sappiamo che $X(t)$ è stazionario in senso lato e ha valor medio nullo. Essendo anche Gaussiano, è pure stazionario in senso stretto. Poiché il sistema è lineare, il processo di uscita sarà a sua volta Gaussiano e anche stazionario, visto che il sistema è anche stazionario. Calcoliamo allora la funzione valor medio e la funzione di autocorrelazione di $Y(t)$. Si trova facilmente che $Y(t)$ ha media nulla:

$$\eta_Y = \eta_X H(0) = 0 \qquad \text{(E8.14.1)}$$

Lo spettro di potenza di $Y(t)$ si calcola quasi altrettanto facilmente:

$$S_Y(f) = S_X(f)|H(f)|^2 = \frac{N_0}{2} \cdot \frac{1}{1+(2\pi f\alpha)^2} \qquad \text{(E8.14.2)}$$

così come la funzione di autocorrelazione:

$$R_Y(\tau) = \mathcal{F}^{-1}\left[S_Y(f)\right] = \frac{N_0}{4\alpha}\exp\left(-\frac{|\tau|}{\alpha}\right) \qquad \text{(E8.14.3)}$$

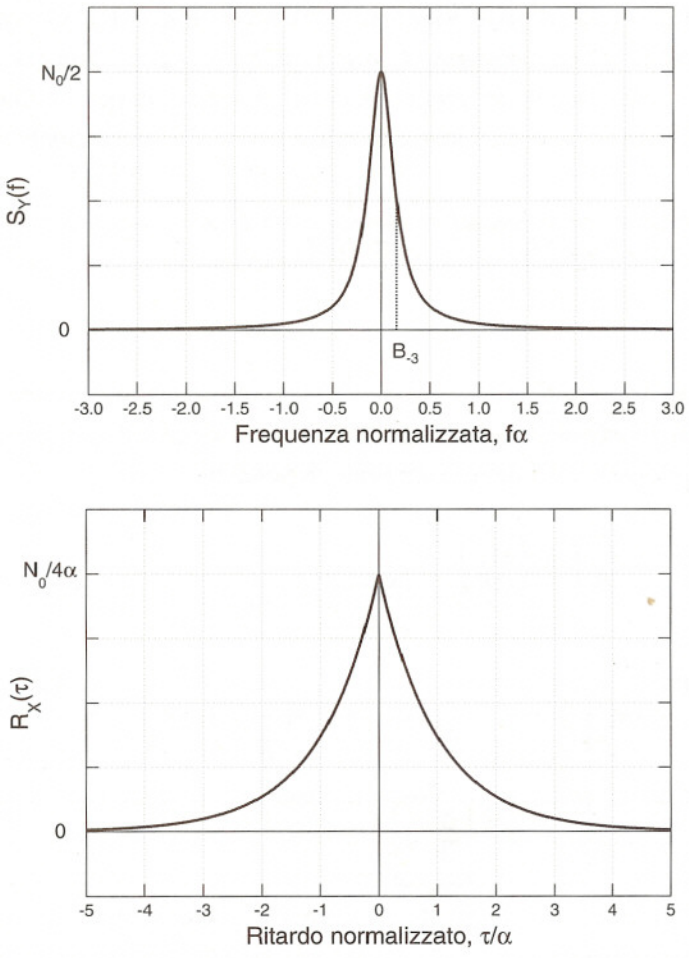


Figura 8.33 Densità spettrale di potenza (a) e funzione di autocorrelazione (b) del processo di Ornstein-Uhlenbeck di Figura 8.32

La densità di potenza di $Y(t)$ (rappresentata in Figura 8.33a) viene comunemente detta *Lorentziana*. Inoltre, un processo Gaussiano con la funzione di autocorrelazione (E8.14.3) (visibile in Figura 8.33b) è chiamato *processo di Ornstein-Uhlenbeck*.

È facile trovare la potenza del segnale di uscita:

$$P_Y = R_Y(0) = \frac{N_0}{4\alpha} = \frac{\pi N_0 \cdot B_{-3}}{2} \qquad \text{(E8.14.4)}$$

che ovviamente cresce al crescere della *banda* a –3 dB (B_{-3}) della squadra *R-C* usata come filtro passa-basso. Siccome $Y(t)$ ha media nulla, la potenza del processo uguaglia la sua varianza: $\sigma_Y^2 = P_Y = \pi N_0 \cdot B_{-3}/2$. La densità del primo ordine di $Y(t)$ *non dipende dal tempo* perché il processo, in quanto Gaussiano e stazionario in senso lato, è stazionario in senso stretto. In conclusione:

$$f_Y(y;t) \equiv f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} e^{-\frac{y^2}{2\sigma_Y^2}} = \frac{1}{\sqrt{\pi^2 N_0 B_{-3}}} e^{-\frac{y^2}{\pi N_0 B_{-3}}} \qquad \text{(E8.14.5)}$$

❑

Esempio 8.15

Consideriamo lo schema della Figura 8.34, nel quale il processo $N(t)$ è Gaussiano stazionario e ha densità spettrale di potenza

$$S_N(f) = N_0\left(1 - \frac{|f|}{B}\right)\text{rect}\left(\frac{f}{2B}\right) \qquad \text{(E8.15.1)}$$

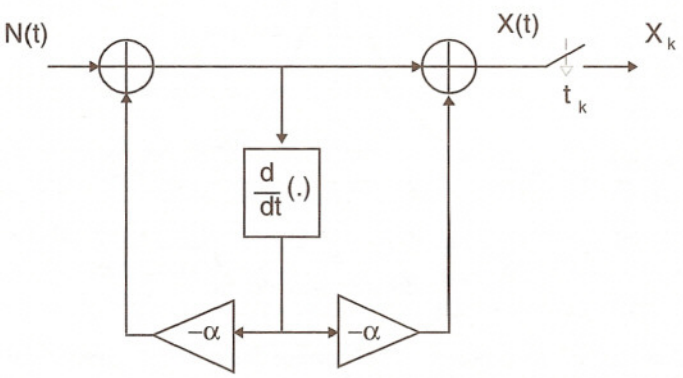


Figura 8.34 Sistema dell'Esempio 8.15

Campionando il processo $X(t)$ agli istanti $t_k = k/B$, con $k = 1,2,\ldots,n$ si ottiene un insieme di n variabili aleatorie $X_k = X(t_k)$, con $k = 1,2,\ldots,n$. Determiniamo la densità di probabilità congiunta $f_{\mathbf{X}}(x_1, x_2, .., x_N)$ di tali variabili aleatorie.

Cominciamo calcolando la funzione di trasferimento $H(s)$ del SLS di Figura 8.43. Sfruttando l'analogia di questo schema con la realizzazione canonica di un sistema a tempo discreto (Figura 6.36), si trova immediatamente

$$H(s)=\frac{1-s\alpha}{1+s\alpha} \tag{E8.15.2}$$

Quindi la risposta in frequenza $H(f)$ del sistema è

$$H(f)=\frac{1-j2\pi f\alpha}{1+j2\pi f\alpha} \tag{E8.15.3}$$

Il processo d'ingresso $N(t)$ è per ipotesi stazionario. Antitrasformando la funzione densità spettrale di potenza se ne ricava la funzione di autocorrelazione:

$$R_N(\tau)=\mathcal{F}^{-1}\left[S_N(f)\right]=N_0\,B\,\mathrm{sinc}^2(B\tau) \tag{E8.15.4}$$

Il valor medio di $N(t)$ è nullo perché $\lim_{\tau\to\infty} R_N(\tau)=0$. Inoltre, il processo $X(t)$ è anch'esso Gaussiano e stazionario perché è l'uscita di un SLS con un processo Gaussiano stazionario in ingresso. Il valor medio di $X(t)$ è nullo, e la sua densità spettrale di potenza è data da

$$S_X(f)=S_N(f)\left|H(f)\right|^2=S_N(f)\frac{1+(2\pi f\alpha)^2}{1+(2\pi f\alpha)^2}=S_N(f) \tag{E8.15.5}$$

Questo risultato si spiega osservando che la risposta in ampiezza del filtro dato è *ovunque pari a* 1 (filtro passa-tutto)! I processi di uscita e di ingresso sono *statisticamente equivalenti*: essi hanno stesso valor medio e funzione di autocorrelazione; essendo entrambi Gaussiani, questo comporta che hanno *le stesse statistiche di qualunque ordine*. Le variabili aleatorie $(X_1, X_2, \ldots, X_n)$ sono estratte da un processo Gaussiano e sono quindi un sistema di n variabili congiuntamente Gaussiane. Per determinare la densità di probabilità congiunta $f_{\mathbf{X}}(x_1, x_2, .., x_N)$ occorre determinarne il vettore dei valori medi $\boldsymbol{\eta}_{\mathbf{X}}$ e la matrice di covarianza $\mathbf{C}_{\mathbf{X}}$. Le componenti del vettore $\boldsymbol{\eta}_{\mathbf{X}}$ sono tutte *nulle* poiché $X(t)$ è un processo a valor medio nullo. L'elemento generico della matrice $\mathbf{C}_{\mathbf{X}}$ è dato da

$$\begin{aligned} c_{X_iX_k} &= \mathrm{E}\left\{\left(X_i-\eta_{X_i}\right)\left(X_k-\eta_{X_k}\right)\right\}=\mathrm{E}\left\{X_i\,X_k\right\}=\mathrm{E}\left\{X(t_i)\,X(t_k)\right\}=R_X(t_i-t_k) \\ &= R_X\left(\frac{i-k}{B}\right)=N_0\,B\,\mathrm{sinc}^2(i-k)=N_0\,B\,\delta[i-k] \end{aligned} \tag{E8.15.6}$$

La matrice è allora *diagonale* perché le variabili sono *a due a due incorrelate*.

Essendo congiuntamente Gaussiane, esse sono anche *indipendenti*. Inoltre, tutte le variabili (a valor medio nullo) hanno stessa varianza $\sigma_X^2 = N_0 B$. La densità di probabilità congiunta $f_{\mathbf{X}}(x_1, x_2, .., x_N)$ è allora semplificabile come segue:

$$f_{\mathbf{X}}(x_1, x_2, .., x_N) = \prod_{i=1}^{n} f_X(x_i) \tag{E8.15.7}$$

ove $f_X(x)$, la densità di probabilità di ciascuna delle variabili $(X_0, X_1, \ldots, X_{n-1})$, è data da

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{x^2}{2\sigma_X^2}} \tag{E8.15.8}$$

In conclusione,

$$f_{\mathbf{X}}(x_1, x_2, .., x_N) = \frac{1}{\sqrt{2\pi(N_0 B)^n}} \exp\left(-\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{2N_0 B}\right) \tag{E8.15.9}$$

❑

Esempio 8.16

Consideriamo un processo Gaussiano $X(t)$ caratterizzato dalla funzione valor medio $\eta_X(t) = 0$ e dalla funzione di autocorrelazione

$$R_X(t_1, t_2) = N_0[\delta(t_1 - t_2) + \delta(t_1 + t_2)] \tag{E8.16.1}$$

Applichiamo questo processo, che risulta evidentemente *non stazionario*, al sistema lineare stazionario della Figura 8.35 (con $\alpha = L/R$) e determiniamo la densità di probabilità del primo ordine $f_Y(y;t)$ del processo di uscita $Y(t)$.

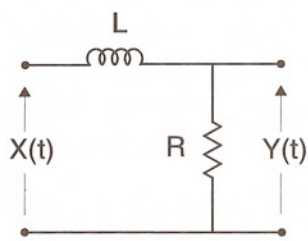


Figura 8.35 Sistema dell'Esempio 8.16

Il processo $Y(t)$ è *Gaussiano* essendo generato da un sistema *lineare* avente in

ingresso un processo $X(t)$ Gaussiano. La densità di probabilità $f_Y(y;t)$ è quindi del tipo

$$f_Y(y;t) = \frac{1}{\sqrt{2\pi}\sigma_Y(t)} e^{-\frac{[y-\eta_Y(t)]^2}{2\sigma_Y^2(t)}} \qquad \text{(E8.16.2)}$$

ove $\eta_Y(t)$ e $\sigma_Y^2(t)$ sono rispettivamente le funzioni valor medio e varianza di $Y(t)$. Ricordiamo che per un processo non stazionario la funzione $\sigma_Y^2(t)$ è pari a

$$\sigma_Y^2(t) = E\left\{\left[Y(t) - \eta_Y(t)\right]^2\right\} = R_Y(t,t) - \eta_Y^2(t) \qquad \text{(E8.16.3)}$$

La funzione $\eta_Y(t)$ si può ricavare dalla relazione (8.4.6):

$$\eta_Y(t) = \eta_X(t) \otimes h(t) \qquad \text{(E8.16.4)}$$

ove $h(t)$, la risposta impulsiva del SLS, nel nostro caso è

$$h(t) = \frac{1}{\alpha} \exp\left(-\frac{t}{\alpha}\right) \mathrm{u}(t) \qquad \text{(E8.16.5)}$$

Poiché per ipotesi $\eta_X(t) = 0$, anche la funzione $\eta_Y(t)$ è identicamente nulla. Per determinare la funzione $\sigma_Y^2(t)$ calcoliamo poi la funzione $R_Y(t_1,t_2)$ attraverso la (8.4.10):

$$R_Y(t_1,t_2) = R_X(t_1,t_2) \otimes h(t_1) \otimes h(t_2) \qquad \text{(E8.16.6)}$$

Sostituendo, si ricava

$$R_Y(t_1,t_2) = N_0\left[\delta(t_1 - t_2) + \delta(t_1 + t_2)\right] \otimes h(t_1) \otimes h(t_2)$$

Svolgendo per prima l'operazione di convoluzione rispetto alla variabile t_1 (e considerando t_2 come una costante) si ha:

$$R_Y(t_1,t_2) = N_0\left[h(t_1 - t_2) + h(t_1 + t_2)\right] \otimes h(t_2) \qquad \text{(E8.16.7)}$$

Scrivendo in modo esplicito la convoluzione rispetto alla variabile t_2, si ottiene:

$$R_Y(t_1,t_2) = N_0 \int_{-\infty}^{+\infty} \left[h(t_1 - \gamma) + h(t_1 + \gamma)\right] h(t_2 - \gamma)\, d\gamma \qquad \text{(E8.16.8)}$$

dalla quale si ricava infine:

$$\sigma_Y^2(t) = R_Y(t,t) = N_0 \int_{-\infty}^{+\infty} \left[h(t-\gamma) + h(t+\gamma)\right] h(t-\gamma)\, d\gamma$$

$$= N_0 \int_{-\infty}^{+\infty} \left[h(\beta) + h(2t-\beta)\right] h(\beta)\, d\beta$$

$$= \frac{N_0}{\alpha^2} \int_{-\infty}^{+\infty} \exp\left(-\frac{2\beta}{\alpha}\right) \mathrm{u}(\beta)\, d\beta + \frac{N_0}{\alpha^2} \int_{-\infty}^{+\infty} \exp\left(-\frac{2t-\beta+\beta}{\alpha}\right) \mathrm{u}(\beta)\,\mathrm{u}(2t-\beta)\, d\beta$$

$$= \frac{N_0}{\alpha^2} \int_{0}^{+\infty} \exp\left(-\frac{2\beta}{\alpha}\right) d\beta + \frac{N_0}{\alpha^2} \exp\left(-\frac{2t}{\alpha}\right) \int_0^{2t} 1\, d\beta \cdot \mathrm{u}(t)$$

$$= \frac{N_0}{2\alpha} + \frac{2N_0}{\alpha^2}\, t \exp\left(-\frac{2t}{\alpha}\right) \mathrm{u}(t) \qquad \text{(E8.16.9)}$$

La varianza non è costante perché il processo $Y(t)$ non è stazionario. ❑

Esempio 8.17

Troviamo la funzione densità di probabilità $f_Z(z)$ della variabile aleatoria Z in Figura 8.36, nell'ipotesi in cui il processo $N(t)$ sia Gaussiano bianco con densità spettrale di potenza $S_N(f) = 2T$.

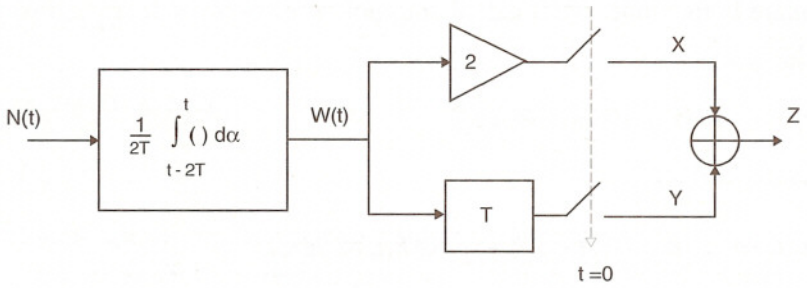


Figura 8.36 Sistema dell'Esempio 8.17

Dallo schema della figura, si vede che la variabile aleatoria Z risulta espressa dalla relazione

$$Z = X + Y = 2W(0) + W(-T) \qquad \text{(E8.17.1)}$$

Il processo $W(t)$ è *Gaussiano e stazionario* perché è la risposta di un SLS a un processo Gaussiano stazionario. Il valor medio di $N(t)$ è *nullo* (processo bianco), e quindi sarà identicamente nulla anche la funzione valor medio di $W(t)$. Osserviamo che quest'ultimo viene generato da un *integratore a finestra*

mobile la cui risposta impulsiva è

$$h(t) = \frac{1}{2T}\text{rect}\left(\frac{t-T}{2T}\right) \tag{E8.17.2}$$

La funzione di autocorrelazione $R_W(\tau)$ di $W(t)$ è allora data da (si veda anche la (E8.14.3)):

$$R_W(\tau) = R_N(\tau) \otimes h(\tau) \otimes h(-\tau) = \left(1 - \frac{|\tau|}{2T}\right)\text{rect}\left(\frac{\tau}{4T}\right) \tag{E8.17.3}$$

Osserviamo poi che le variabili aleatorie $W(0)$ e $W(-T)$, essendo estratte da un processo Gaussiano $W(t)$, sono *congiuntamente Gaussiane* e, quindi, la variabile aleatoria Z è *Gaussiana* perché è una combinazione lineare di variabili congiuntamente Gaussiane. Il valor medio η_Z di Z è pari a

$$\eta_Z = E\{Z\} = 2E\{W(0)\} + E\{W(-T)\} = 2\eta_W(0) + \eta_W(-T) = 0 \tag{E8.17.4}$$

mentre la sua varianza σ_Z^2 è

$$\begin{aligned}\sigma_Z^2 &= E\left\{(Z-\eta_Z)^2\right\} = E\left\{Z^2\right\} = E\left\{\left[2W(0) + W(-T)\right]^2\right\} \\ &= 4E\left\{W(0)^2\right\} + E\left\{W(-T)^2\right\} + 4E\left\{W(0)W(-T)\right\} \\ &= 4R_W(0) + R_W(0) + 4R_W(T) = 4 + 1 + 4 \cdot 1/2 = 7\end{aligned} \tag{E8.17.5}$$

In conclusione, la densità di probabilità $f_Z(z)$ della variabile aleatoria Z è

$$f_Z(z) = \frac{1}{\sqrt{14\pi}} e^{-\frac{z^2}{14}} \tag{E8.17.6}$$

❑

8.7 Processi ergodici

8.7.1 Ergodicità del valore medio

Abbiamo intenzionalmente ritardato fino a questo paragrafo la discussione di un aspetto molto importante sia dal punto di vista concettuale sia da quello pratico. Come sappiamo, anche nella definizione delle grandezze statistiche *semplificate* (funzione valor medio, funzione di autocorrelazione) di un processo aleatorio è sempre prevista un'operazione di *valor medio* relativamente a tutte le possibili funzioni campione del processo. Abbiamo chiamato quest'operazione valor medio *statistico*, o, in alternativa, valor medio *sull'insieme* (sottinteso: delle fun-

zioni campione). Dal punto di vista teorico, quest'operazione non comporta alcuna difficoltà, supponendo beninteso di conoscere le opportune funzioni densità di probabilità del processo che consentono di effettuare le operazioni di media previste (densità del primo ordine per la funzione valor medio, del secondo ordine per la funzione di autocorrelazione ecc.). In pratica, le funzioni densità di un processo non sono generalmente note, e in molti casi esse non sono neanche ricavabili con ragionevole accuratezza attraverso delle misure. Per fare misure statistiche significative su di un processo dovremmo infatti osservare un numero significativo di funzioni campione rispetto a tutte quelle possibili. Molto spesso, di un dato processo viene invece osservata solo *una* realizzazione, quella cioè che si è verificata in una certa prova dell'esperimento. Sorge allora spontanea la domanda: è possibile riuscire a effettuare misure su questa *particolare* realizzazione e in qualche modo ricavare il comportamento statistico *generale* del processo? La risposta a questa domanda ruota attorno alla nozione di *ergodicità* di un processo, che dobbiamo discutere in un certo dettaglio[2].

Supponiamo di sapere con certezza che il processo $X(t)$ sotto osservazione è *stazionario* in media, e poniamoci il problema di stimarne il *valor medio* η_X (costante). Ciò che si ha a disposizione in una prova dell'esperimento è una certa *realizzazione* del processo $x(t;\overline{\omega})$ corrispondente al particolare risultato $\overline{\omega}$ dell'esperimento stesso. Di questa certa realizzazione che, una volta osservata, è un segnale *determinato*, possiamo certamente calcolare il valor medio *temporale* (1.3.7*a*)

$$x_m(\overline{\omega}) \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(\overline{\omega};t)\,dt \tag{8.7.1}$$

e possiamo chiederci poi quale sia la relazione di questa quantità con il valor medio *statistico* (o *sull'insieme*) η_X. Due questioni si pongono da questo punto di vista: i) in generale, il valor medio temporale *dipende dalla particolare*

---

[2] La parola *ergodico* (dal greco *ergodes*, laborioso) appare per la prima volta in meccanica statistica con un'accezione molto simile a quella usata qui. In meccanica statistica, infatti, le proprietà macroscopiche di un insieme di particelle microscopiche (ad esempio, la pressione di un gas) vengono ricavate attraverso una media del comportamento *sull'insieme* delle particelle. La proprietà di ergodicità del sistema dice però che le stesse proprietà macroscopiche possono essere anche ricavate seguendo l'evoluzione *temporale* di una *singola* particella che è *tipica* del comportamento globale dell'insieme.

funzione campione e quindi dal particolare risultato dell'esperimento, come suggerisce la notazione nella (8.7.1). Non è detto quindi che, osservando più volte il processo, si ottenga sempre uno stesso valor medio temporale dalle varie realizzazioni; ii) anche se il valore è sempre lo stesso, non è detto che questo coincida con il valor medio statistico η_X.

Alcuni processi sono invece tali che i) tutte le funzioni campione hanno lo stesso valor medio temporale, e ii) quest'ultimo coincide col valor medio statistico η_X: sono i processi *ergodici in media*. Per quanto riguarda il valor medio, ogni funzione campione di un processo ergodico in media "si comporta" come il processo nella sua globalità: il valor medio temporale (di ogni funzione campione) è uguale al valor medio statistico, come rappresentato sinteticamente nella Figura 8.37. Per il processo ergodico, una grandezza *statistica* può essere dunque ricavata attraverso una media *temporale* su una *qualunque* realizzazione del processo stesso. È chiaro che un processo *non stazionario* in media non può essere ergodico. Infatti, se il valor medio statistico non è *costante*, non ha senso porsi il problema dell'uguaglianza tra questo e un valor medio *temporale* che è un singolo valore per definizione.

Sotto quali condizioni un processo è ergodico in media? Per rispondere a questa domanda, osserviamo che in generale il valor medio temporale delle funzioni campione è una quantità numerica che *dipende dal risultato dell'esperimento*: in sintesi, è una *variabile aleatoria*. Con la notazione più appropriata, la (8.7.1) diventa

$$X_m \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t)dt \tag{8.7.2}$$

ove abbiamo omesso, come di consueto, la dipendenza dal risultato dell'esperimento. Le due condizioni sotto cui il processo è ergodico si possono dunque riassumere dicendo che il *valore atteso* di questa variabile aleatoria deve essere uguale al valor medio del processo η_X, e la *varianza* della stessa variabile deve essere *nulla*, dimodoché tutte le funzioni campione presentano sempre (cioè con probabilità 1) un valor medio temporale pari a η_X stesso!

In pratica, non potremo osservare il processo (o una sua funzione campione) su di un intervallo *illimitato*. Cominciamo allora a considerare la variabile aleatoria

$$X_T \triangleq \frac{1}{T} \int_{-T/2}^{T/2} X(t)dt \tag{8.7.3}$$

ottenuta mediando le funzioni campione sull'intervallo *limitato* $[-T/2,-T/2]$. Evidentemente,

$$X_m = \lim_{T\to\infty} X_T \quad , \quad \eta_{X_m} = \lim_{T\to\infty} \eta_{X_T} \quad , \quad \sigma^2_{X_m} = \lim_{T\to\infty} \sigma^2_{X_T} \tag{8.7.4}$$

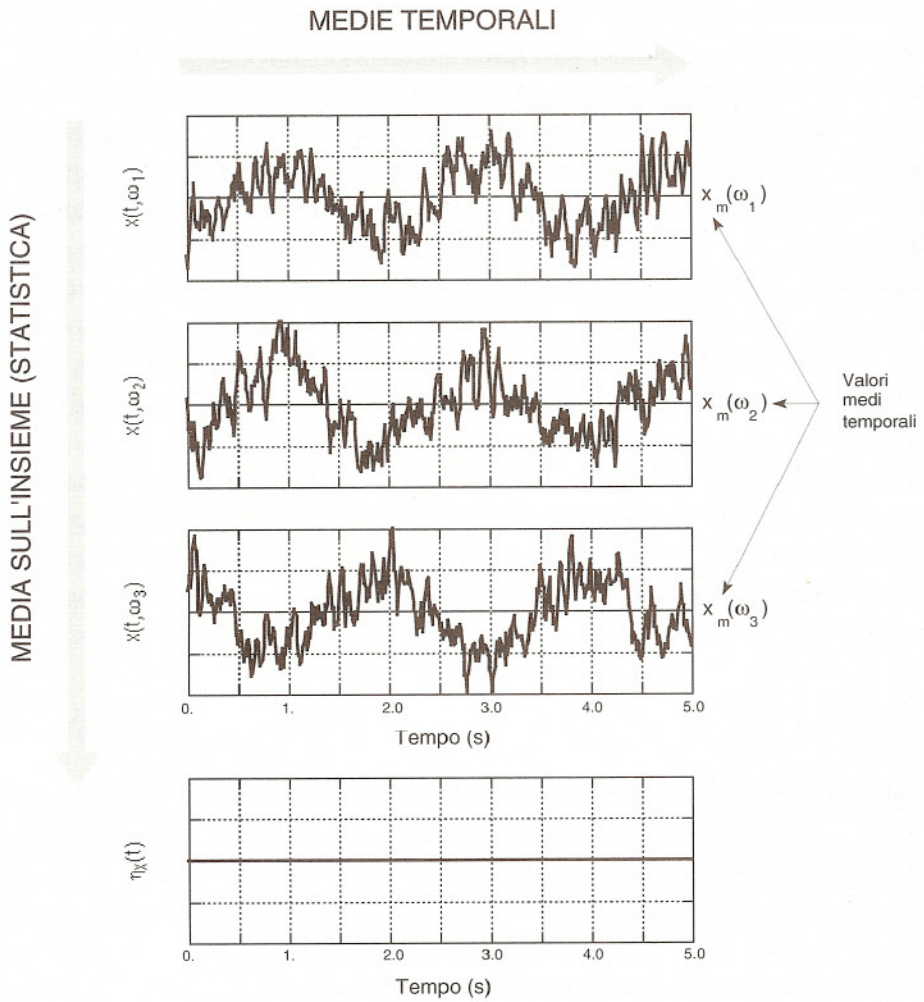


Figura 8.37 Rappresentazione sintetica dell'ergodicità del valor medio

La (8.7.3) suggerisce anche un modo per ottenere *in pratica* le medie temporali richieste. Il valore X_T si può interpretare infatti come *l'uscita di un integratore a finestra mobile* su di un intervallo temporale di ampiezza T, valutata a un certo

istante, come suggerisce la Figura 8.38. In figura è anche rappresentata la risposta impulsiva di questo filtro:

$$h(t) = \frac{1}{T}\,\mathrm{rect}\left(\frac{t - T/2}{T}\right) \tag{8.7.5}$$

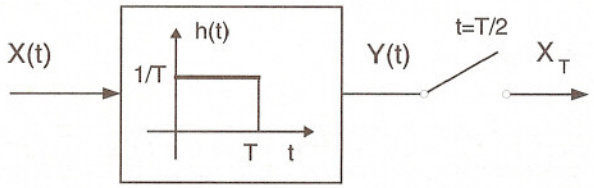


Figura 8.38 Variabile aleatoria valor medio temporale delle funzioni campione di un processo

L'integratore a finestra mobile con un intervallo di osservazione molto ampio (ricordiamo che in teoria $T \to \infty$) è un filtro passa-basso a banda molto stretta. La risposta in ampiezza di tale filtro è infatti $H(f) = |\mathrm{sinc}(fT)|$ il cui lobo principale si estende da $-1/2T$ a $1/2T$. La media temporale di una realizzazione del processo si può dunque eseguire valutando l'uscita di un filtro passa-basso a banda stretta a un opportuno istante (in pratica, dopo che i fenomeni transitori in uscita al filtro si sono estinti).

Torniamo dunque al sistema di Figura 8.38. Poiché il processo di ingresso è stazionario in senso lato, il valor medio dell'uscita di tale filtro è *costante* (il processo $Y(t)$ è ancora stazionario) e vale:

$$\eta_{X_T} = \eta_Y(T/2) \equiv \eta_Y = \eta_X H(0) = \eta_X \tag{8.7.6}$$

e quindi

$$\eta_{X_m} = \lim_{T\to\infty} \eta_{X_T} = \eta_X \tag{8.7.7}$$

Questo risultato *non* dipende da alcuna particolare ipotesi sul processo $X(t)$, e quindi la prima condizione per l'ergodicità è automaticamente verificata per ogni processo stazionario (in senso lato). Dobbiamo però calcolare la *varianza* $\sigma^2_{X_T}$. A questo proposito, torna ancora utile la Figura 8.38. Sapendo che $Y(t)$ è stazionario, si può infatti scrivere che

$$\sigma^2_{X_T} = \sigma^2_Y = C_Y(0) \tag{8.7.8}$$

dove $C_Y(\tau)$ è la funzione di autocovarianza del processo (stazionario) $Y(t)$.

D'altronde:

$$C_Y(0) = C_X(\tau) \otimes h(t) \otimes h(-\tau)\big|_{\tau=0} = C_X(\tau) \otimes \frac{1}{T}\left(1-\frac{|\tau|}{T}\right)\text{rect}\left(\frac{\tau}{2T}\right)\Bigg|_{\tau=0}$$

$$= \int_{-\infty}^{\infty} C_X(\alpha)\frac{1}{T}\left(1-\frac{|\tau-\alpha|}{T}\right)\text{rect}\left(\frac{\tau-\alpha}{2T}\right)d\alpha\Bigg|_{\tau=0} \tag{8.7.9}$$

per cui si conclude che

$$\sigma^2_{X_T} = \frac{1}{T}\int_{-\infty}^{\infty} C_X(\alpha)\left(1-\frac{|\alpha|}{T}\right)\text{rect}\left(\frac{\alpha}{2T}\right)d\alpha = \frac{1}{T}\int_{-T}^{T} C_X(\tau)\left(1-\frac{|\tau|}{T}\right)d\tau \tag{8.7.10}$$

La condizione di *ergodicità del valor medio* è presto trovata: la varianza $\sigma^2_{X_m}$ di X_m deve essere *nulla* e quindi deve valere:

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T}^{T} C_X(\tau)\left(1-\frac{|\tau|}{T}\right)d\tau = 0 \tag{8.7.11}$$

L'ergodicità del valor medio, che è una statistica di ordine 1 del processo, è subordinata al verificarsi di una condizione su di una statistica di ordine 2, cioè di ordine superiore.

8.7.2 Ergodicità della funzione di autocorrelazione - Ergodicità in senso stretto

Quando si affrontano questioni di *ergodicità* dei processi aleatori, è utile definire un *operatore* valor medio temporale $\langle\cdot\rangle$ che si può applicare a un qualunque segnale *determinato* $x(t)$:

$$\langle x(t)\rangle \triangleq \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x(t)dt = x_m \tag{8.7.12}$$

La proprietà di ergodicità del valor medio vista nel paragrafo precedente si può allora esprimere in modo sintetico come segue:

$$E\{X(t)\} = \langle x(\omega;t)\rangle \tag{8.7.13a}$$

dove $x(t;\omega)$ è la *generica* funzione campione del processo $X(t)$. Con notazione impropria, si può riassumere la proprietà di ergodicità del valor medio

$$E\{X(t)\} = \langle X(t)\rangle \tag{8.7.13b}$$

intendendo in realtà la (8.7.13*a*). Spesso, la proprietà di ergodicità viene

riassunta nell'uguaglianza (un po' semplicistica) "*valor medio statistico=valor medio temporale*".

L'operatore valor medio temporale torna utile anche per esprimere la funzione di autocorrelazione di un segnale *determinato* a potenza finita (4.4.31):

$$R_x(\tau) = \langle x(t)x(t-\tau)\rangle \triangleq \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t-\tau)dt \tag{8.7.14}$$

È abbastanza chiaro allora come estendere la proprietà di ergodicità anche alla funzione di autocorrelazione. Un processo si dice *ergodico in autocorrelazione* se, indipendentemente dalla particolare funzione campione, risulta:

$$R_X(\tau) \triangleq E\{X(t)X(t-\tau)\} = \langle x(\omega;t)x(\omega;t-\tau)\rangle = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x(\omega;t)x(\omega;t-\tau)dt \tag{8.7.15}$$

cioè se la funzione di autocorrelazione *statistica* del processo è uguale alla funzione di autocorrelazione *temporale* di una qualunque realizzazione. È chiaro che un processo non stazionario in senso lato *non* può essere ergodico in autocorrelazione in quanto la funzione di autocorrelazione calcolata come media *temporale* dipende soltanto dalla variabile τ e non da t, e non può quindi uguagliare la funzione di autocorrelazione statistica di un processo non stazionario.

L'ergodicità in autocorrelazione è importante perché se è possibile ricavare l'autocorrelazione del processo da una singola funzione campione, allora da questa è anche possibile ricavare (con semplice trasformata di Fourier) la *densità spettrale* del processo stesso. In questo caso, il processo può essere analizzato con le medesime tecniche (di tipo *temporale* e non *statistico*) usate per l'analisi dei segnali determinati. Le condizioni per l'ergodicità in autocorrelazione (su cui non insisteremo per semplicità) coinvolgono grandezze statistiche del *quarto* ordine del processo, e sono di difficile verifica, salvo casi particolari (i processi Gaussiani).

Un processo ergodico in valor medio e funzione di autocorrelazione si dice *ergodico in senso lato*. Se però la proprietà di ergodicità vale per una *qualunque* grandezza statistica estratta dal processo, allora il processo è *ergodico in senso stretto* e gode della proprietà seguente:

$$E\{g[X(t), X(t-\tau_1), \ldots, X(t-\tau_{n-1})]\} = \langle g[x(\omega;t), x(\omega;t-\tau_1), \ldots, x(\omega;t-\tau_{n-1})]\rangle \tag{8.7.16}$$

dove $g[\cdot]$ è una *qualunque* funzione reale di n variabili e dove evidentemente il valor medio temporale *non* dipende da ω. Si noti che nell'enunciare questa proprietà è stata implicitamente assunta la stazionarietà in senso stretto del processo.

Esempio 8.18

Discutiamo l'ergodicità in valor medio del processo parametrico

$$X(t) = A \tag{E8.18.1}$$

dove A è una variabile aleatoria con densità di probabilità $f_A(a)$ assegnata. Il problema è ben posto, perché il processo è stazionario in media:

$$E\{X(t)\} = E\{A\} = \eta_A \tag{E8.18.2}$$

Nell'Esempio 8.4 abbiamo perfino dimostrato la stazionarietà *in senso stretto* di questo processo.

Le funzioni campione di $X(t)$, rappresentate in Figura 8.10, sono delle *costanti*. Consideriamo una qualunque di queste funzioni: quella, ad esempio, corrispondente al valore $\bar{a}$ della variabile A:

$$x(t;\bar{\omega}) = \bar{a} \tag{E8.18.3}$$

Il valor medio *temporale* di questa funzione campione è chiaramente $\bar{a}$, che in generale è *diverso* dal valor medio statistico η_A di $X(t)$: il processo *non* è ergodico in media. Il lettore può dimostrare facilmente che la condizione (8.7.11) per l'ergodicità in media *non* è verificata. ❑

Esempio 8.19

Dimostriamo che il processo aleatorio

$$X(t) = a\cos(2\pi f_0 t + \Theta) \tag{E8.19.1}$$

ove a ed f_0 sono noti e $\Theta \in \mathcal{U}(0,2\pi)$ è *ergodico in senso lato*. Nell'Esempio 8.5 abbiamo già dimostrato che il processo è *stazionario* in senso lato, e quindi il problema è ben posto. Nello stesso esempio, abbiamo trovato

$$\eta_X = 0 \quad , \quad R_X(\tau) = \frac{a^2}{2}\cos(2\pi f_0 \tau) \tag{E8.19.2}$$

CENTRO BIBLIOTECARIO INGEGNERIA PISA

Calcoliamo ora le corrispondenti *medie temporali*:

$$\langle x(\omega;t)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} a\cos(2\pi f_0 t+\Theta)dt = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} a\cos(2\pi f_0 t+\Theta)dt \qquad \text{(E8.19.3)}$$

Sappiamo infatti che la media temporale per un segnale periodico (come *tutte* le funzioni campione del processo) può essere valutata su un singolo periodo anziché su tutto l'asse temporale. Proseguendo il calcolo si trova

$$\langle x(\omega;t)\rangle = \frac{1}{T_0}\left.\frac{a\cos(2\pi f_0 t+\Theta)}{2\pi f_0}\right|_{-T_0/2}^{T_0/2} = 0 \qquad \text{(E8.19.4)}$$

indipendentemente dal valore di Θ, e cioè dalla particolare funzione campione. Il processo è quindi ergodico in media.

Per la funzione di autocorrelazione si ha poi:

$$\begin{aligned}\langle x(\omega;t)x(\omega;t-\tau)\rangle &= \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} a\cos(2\pi f_0 t+\Theta)\,a\cos(2\pi f_0(t-\tau)+\Theta)dt\\ &= \frac{a^2}{2T_0}\int_{-T_0/2}^{T_0/2}\cos(2\pi f_0\tau)dt + \frac{a^2}{2T_0}\int_{-T_0/2}^{T_0/2}\cos(2\pi f_0(2t-\tau)+2\Theta)dt\\ &= \frac{a^2}{2}\cos(2\pi f_0\tau) = R_X(\tau) \qquad \text{(E8.19.5)}\end{aligned}$$

Il processo è anche ergodico in autocorrelazione e quindi ergodico in senso lato.

❑

Sommario

In questo capitolo abbiamo introdotto e discusso i principali concetti necessari allo studio dei segnali aleatori. Un *processo aleatorio* $X(t)$ (cioè la modellizzazione matematica di un segnale non noto a priori) è una *corrispondenza* che associa a ogni risultato di un esperimento aleatorio una *funzione* campione selezionata in una certa famiglia. Se si fissa un istante temporale, i vari valori che le funzioni campione possono assumere definiscono una variabile aleatoria. La caratterizzazione del processo è allora *completa* quando è nota la densità di probabilità di ordine n del processo, con n qualunque, cioè la densità congiunta delle n variabili aleatorie $X(t_1)$, $X(t_2),\ldots, X(t_n)$ estratte dal processo a istanti $(t_1,t_2\ldots,t_n)$ arbitrari. Poiché la caratterizzazione completa è assi complicata, ci si accontenta spesso dei parametri statistici semplificati: la *funzione valor medio*

$\eta_X(t) \triangleq E\{X(t)\}$ e la *funzione di autocorrelazione* (oppure quella di autocovarianza) $R_X(t_1,t_2) \triangleq E\{X(t_1)X(t_2)\}$ ($C_X(t_1,t_2) \triangleq E\{[X(t_1)-\eta_X(t_1)][X(t_2)-\eta_X(t_2)]\}$).

Una proprietà importante dei processi aleatori è la *stazionarietà*. Un processo è *stazionario in senso stretto* se tutte le sue grandezze statistiche non variano cambiando l'origine del riferimento temporale (sono cioè *invarianti* alle traslazioni temporali). Questo avviene se e solo se la stessa proprietà di invarianza temporale è posseduta dalla densità di probabilità del processo di qualunque ordine: $f_X(x_1,x_2,...,x_n;t_1+\Delta t,t_2+\Delta t,...,t_n+\Delta t) = f_X(x_1,x_2,...,x_n;t_1,t_2,...,t_n)$ per ogni traslazione Δt. Un tipo di stazionarietà di più semplice verifica è quella *in senso lato*, che richiede l'invarianza temporale delle sole funzioni valor medio e autocorrelazione: $\eta_X(t) = \eta_X$ e $R_X(t_1,t_2) = R_X(t_1-t_2)$. Indicando diversamente gli istanti t_1 e t_2, quest'ultima proprietà si riscrive: $R_X(t,t-\tau) = R_X(\tau)$, ove τ è la distanza temporale tra t_1 e t_2. La funzione di autocorrelazione di un processo stazionario gode delle stesse proprietà presentate dalla funzione di autocorrelazione per i segnali determinati già introdotta nel Capitolo 4.

Quando un processo aleatorio $X(t)$ viene *filtrato* con un SLS avente risposta in frequenza $H(f)$ e risposta impulsiva $h(t)$, non è in generale possibile ricavare la descrizione completa dell'uscita $Y(t)$ a partire da quella dell'ingresso. È però semplice mettere in relazione le statistiche semplificate dei due processi. Nel caso di processo $X(t)$ stazionario (almeno in senso lato) si ha: $\eta_Y = \eta_X H(0)$ e $R_Y(\tau) = R_X(\tau) \otimes h(\tau) \otimes h(-\tau)$.

Si definisce poi la *densità spettrale di potenza* di un processo $X(t)$ stazionario in senso lato come trasformata di Fourier della funzione di autocorrelazione $R_X(\tau)$. Questo implica che lo spettro di potenza di un processo viene modificato dal filtraggio esattamente come lo spettro di potenza di un segnale determinato ($S_Y(f) = S_X(f)|H(f)|^2$) e che ne condivide le stesse proprietà. In particolare, un processo aleatorio stazionario si dice *bianco* quando possiede una densità spettrale di potenza *costante* per tutte le frequenze.

Molti fenomeni naturali possono essere modellati come *processi Gaussiani*. Un processo aleatorio $X(t)$ è *Gaussiano* se n variabili aleatorie da esso estratte agli istanti $(t_1,t_2,...,t_n)$ risultano *congiuntamente Gaussiane* comunque si scelga il valore del parametro intero n e per qualunque n-upla di istanti $(t_1,t_2,...,t_n)$. Questa definizione comporta che il processo è completamente descritto dalle sole funzioni valor medio e autocorrelazione, che è stazionario in senso stretto se lo è in senso lato, e che resta Gaussiano quando viene elaborato con un sistema lineare (non necessariamente stazionario).

La possibilità di ricavare grandezze statistiche di un processo a partire dal-

l'osservazione di *una sola* realizzazione è infine garantita dalla proprietà di *ergodicità*: un processo è ergodico se le grandezze medie statistiche (ad esempio la funzione di autocorrelazione) possono essere ricavate come valor medio temporale della corrispondente grandezza, valutato *su una qualunque* funzione campione del processo stesso. Ciò significa che ogni realizzazione è *tipica* dell'andamento globale del processo, e che non è necessario osservare *tutte* le funzioni campione del processo per ricavare l'andamento delle varie grandezze statistiche.

Esercizi proposti

8.1 La Figura 8.39 mostra il grafico delle cinque funzioni campione che compongono un processo aleatorio $X(t)$. Tenendo conto che queste funzioni sono associate ai cinque risultati di un esperimento *simmetrico* (cioè tutte le funzioni sono *equiprobabili*), dire qual è il particolare valore della funzione distribuzione del secondo ordine $F_X(0,3;2,0)$.

8.2 È assegnato un processo aleatorio $X(t)$ stazionario in senso stretto. Spiegare perché, in generale, il processo $Y(t) = X(t)\mathrm{u}(t)$ non può essere stazionario.

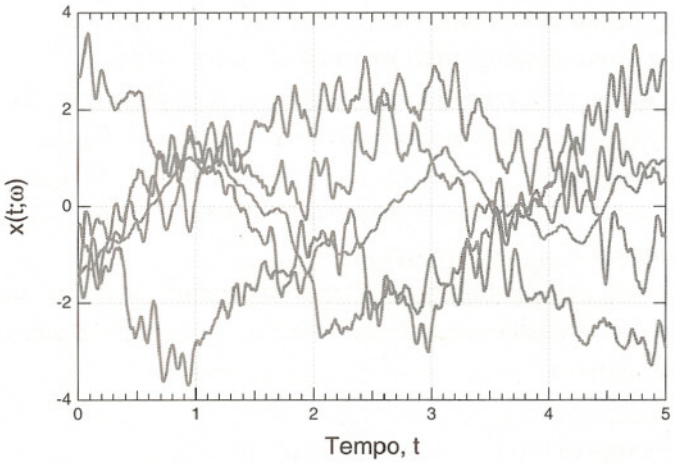


Figura 8.39

8.3 È dato il processo aleatorio

$$X(t) = X_1 \cos(2\pi f_0 t) - X_2 \sin(2\pi f_0 t)$$

ove X_1 e X_2 sono due variabili aleatorie indipendenti Gaussiane a valor

medio nullo e varianza $\sigma_{X_1}^2 = \sigma_{X_2}^2 = 4$. Questo processo è l'ingresso dei due SLS di Figura 8.40 aventi le risposte in frequenza indicate. Ricavare la densità di probabilità del primo ordine $f_Y(y;t)$ del processo $Y(t)$ nei due casi.

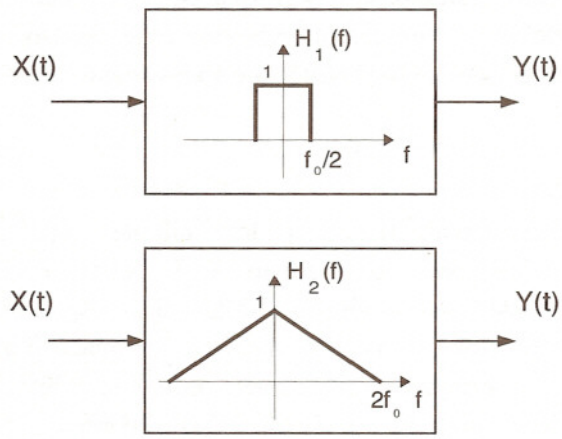


Figura 8.40

8.4 Calcolare e rappresentare la densità spettrale di potenza del processo aleatorio *segnale dati binario stazionario* dell'Esempio 8.9.

8.5 Premessa: *Data una coppia di processi aleatori $Y(t)$ e $Z(t)$ si definisce la funzione di correlazione mutua $R_{YZ}(t,t-\tau) = E\{Y(t)Z(t-\tau)\}$. I processi $Y(t)$ e $Z(t)$ si dicono congiuntamente stazionari se $R_{YZ}(t,t-\tau)$ dipende solo dalla variabile τ, cioè se $R_{YZ}(t,t-\tau) = R_{YZ}(\tau)$. Se i processi sono congiuntamente stazionari è possibile calcolare la densità spettrale di potenza mutua $S_{YZ}(f) = \mathcal{F}[R_{YZ}(\tau)]$.*

Nello schema della Figura 8.41 $X(t)$ è un segnale aleatorio stazionario in senso lato con densità spettrale di potenza $S_X(f) = \xi$ e i due filtri hanno risposta impulsiva

$$h_1(t) = \frac{1}{\alpha}\exp(-\alpha t)\,u(t) \quad , \quad h_2(t) = \frac{1}{\beta}\exp(-\beta t)\,u(t)$$

Dimostrare che i processi $Y_1(t)$ e $Y_2(t)$ sono congiuntamente stazionari e calcolare la funzione di correlazione mutua $R_{Y_1Y_2}(\tau)$ e la densità spettrale di potenza mutua $S_{Y_1Y_2}(f)$. [*Trovare la forma generale di $R_{Y_1Y_2}(\tau)$ e $S_{Y_1Y_2}(f)$ per lo schema della Figura 8.41, prescindendo dal caso particolare, e applicare poi tale risultato al caso particolare indicato.*]

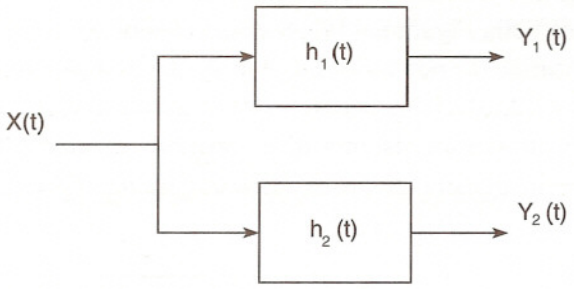


Figura 8.41

8.6 Con la stessa premessa dell'esercizio precedente, calcolare l'espressione generale della *correlazione mutua* $R_{YX}(\tau)$ tra un processo aleatorio stazionario $X(t)$ in ingresso a un SLS e il rispettivo processo di uscita $Y(t)$.

8.7 Una tensione costante v_0 viene disturbata da rumore bianco additivo $N(t)$ avente densità spettrale di potenza $S_N(f) = \zeta$. Per reiettare tale disturbo si usano i due sistemi in cascata mostrati nella Figura 8.42, in cui $B = 1/2T$. Determinare i valori dei coefficienti a, b e c che minimizzano l'*errore quadratico medio*

$$\varepsilon = E\{[Y(t) - v_0]^2\}$$

tra l'uscita $Y(t)$ e il valore costante v_0.

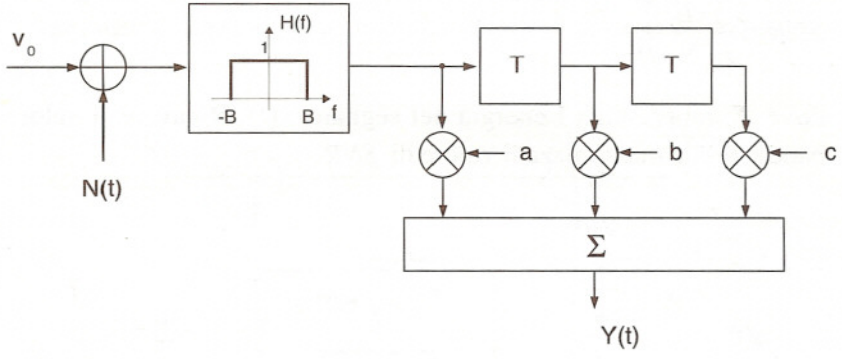


Figura 8.42

8.8 Determinare e rappresentare la densità spettrale di potenza $S_X(f)$ del processo aleatorio $X(t) = 2\sin(2\pi Ft + \Theta)$, ove F è una variabile aleatoria uniformemente distribuita in $[f_0, 2f_0]$ e Θ è una seconda variabile

indipendente da F e uniformemente distribuita in $[-\pi,\pi]$.

8.9 Nello schema della Figura 8.43 il processo aleatorio $N(t)$ è Gaussiano con densità spettrale di potenza $S_N(f)=N_0/2$. Campionando agli istanti $t_k=2kT\,(k=1,2,\ldots,n)$ il processo $X(t)$ in uscita al filtro $H(f)$ (di banda $B=1/T$) si ottiene un sistema di n variabili aleatorie $\{X_k \equiv X(2kT)\}$. Determinare la densità di probabilità congiunta $f_X(x_1,x_2,..,x_n)$ di tale sistema.

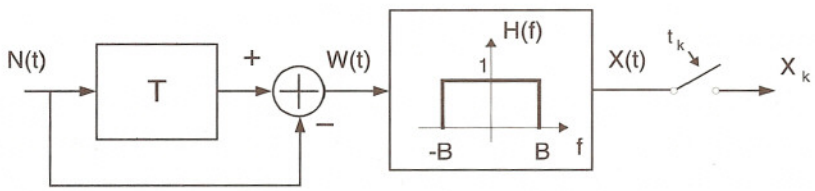


Figura 8.43

8.10 Nello schema di Figura 8.44 il processo di rumore $W(t)$ è stazionario in senso lato e ha densità spettrale di potenza $S_W(f)=N_0/2$. Il segnale $s(t)$ è

$$s(t)=2\sin(2f_0t)\,\mathrm{sinc}^2(f_0t)$$

Si indichi con $N(t)$ il processo di rumore in uscita al filtro di banda B e con $x(t)$ la risposta del filtro a $s(t)$ e si definisca il *rapporto segnale-rumore* in uscita al filtro come

$$SNR=\frac{E_x}{E\{N^2(t)\}}$$

dove E_x rappresenta l'energia del segnale $x(t)$. Ricavare il valore della banda B che massimizza il valore di SNR.

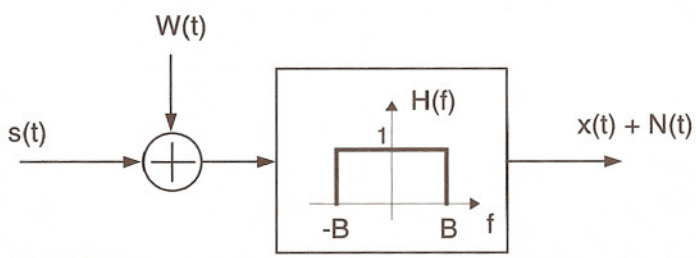


Figura 8.44

8.11 Nella Figura 8.45 il processo SSL $N(t)$ è Gaussiano e ha densità spettrale di potenza $S_N(f) = N_0/2$. Determinare la probabilità che la variabile aleatoria X assuma un valore maggiore di quello della variabile Y.

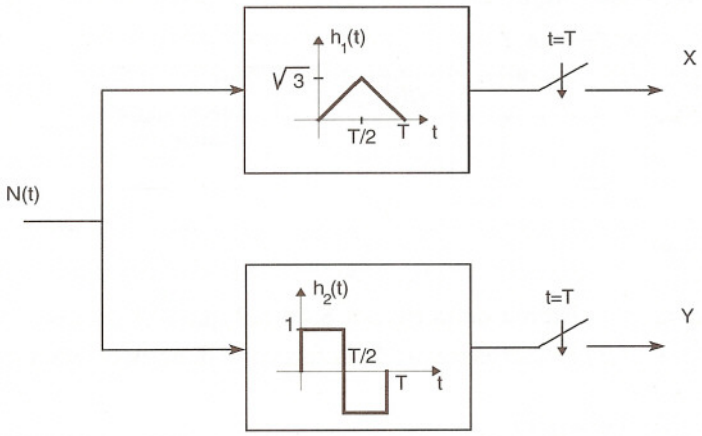


Figura 8.45

8.12 È dato il processo aleatorio Gaussiano stazionario $X(t)$ caratterizzato dalla funzione densità spettrale di potenza

$$S_X(f) = \frac{N_0}{1+(2\pi f\beta)^2}$$

Tale processo è applicato in ingresso al SLS della Figura 8.34 (α, $\beta > 0$). Determinare la densità di probabilità del primo ordine $f_Y(y;t)$ del processo di uscita $Y(t)$.

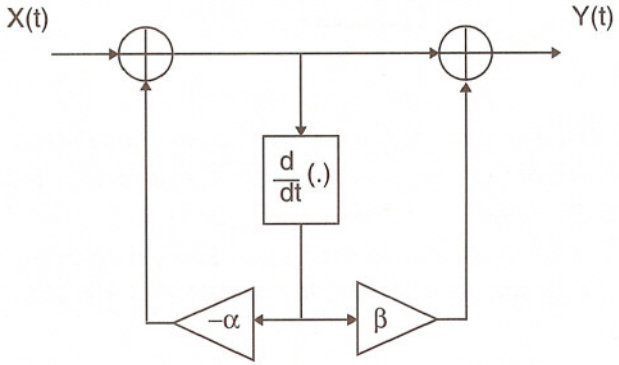


Figura 8.46

8.13 Nello schema della Figura 8.47, il processo $N(t)$ è Gaussiano con densità spettrale di potenza $S_N(f)=\xi$. Dire se il processo di uscita $X(t)$ è Gaussiano e determinarne la funzione valor medio $\eta_X(t)$ e la funzione di autocorrelazione $R_X(t,\tau)$.

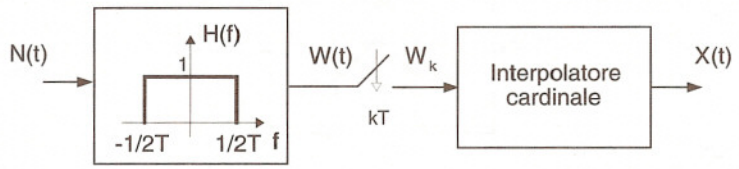


Figura 8.47

8.14 Considerare il sistema della Figura 8.48 nel quale il processo Gaussiano d'ingresso $X(t)$ è caratterizzato dalla funzione di autocorrelazione

$$R_X(\tau)=N_0 B\,\mathrm{sinc}(B\tau)$$

Ricavare la densità di probabilità del primo ordine $f_Y(y;t)$ del processo aleatorio d'uscita $Y(t)$ nei due casi seguenti:

a) $H(f)=1$

b) $H(f)=\mathrm{rect}(f/B)$

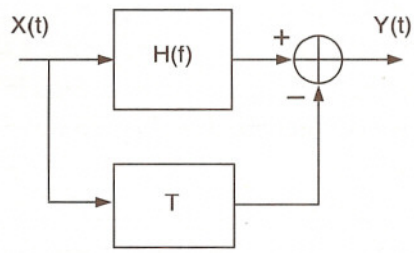


Figura 8.48

8.15 Premessa: *Due processi $X(t)$ e $Y(t)$ si dicono indipendenti se qualunque insieme di variabili aleatorie estratte da $X(t)$ è indipendente da qualunque insieme di variabili aleatorie estratte da $Y(t)$.*

Sono dati due processi stazionari in senso lato e indipendenti $X(t)$ e $Y(t)$ con funzione di autocorrelazione $R_X(\tau)=R_Y(\tau)=\sigma^2\mathrm{sinc}^2(B\tau)$. Costruito il processo

$$Z(t)=X(t)\cos(4\pi Bt)-Y(t)\sin(4\pi Bt)$$

dimostrare che $Z(t)$ è stazionario in senso lato. Determinare e rappresentare poi la densità spettrale di potenza $S_Z(f)$ di $Z(t)$.

8.16 Sono assegnati due processi aleatori $X(t)$ e $Y(t)$ stazionari in senso lato, indipendenti (si veda l'Esercizio 8.15), e aventi le densità spettrali di potenza rispettivamente $S_X(f)$ ed $S_Y(f)$ rappresentate nella Figura 8.49. Trovare (almeno) un valore del parametro K per il quale $Z(t)$ in Figura 8.50 risulta stazionario in senso lato e, con tale valore di K, calcolare la potenza media statistica P_W del processo $W(t)$ in uscita al filtro della Figura 8.51.

8.17 Dimostrare che condizione sufficiente per l'ergodicità del valor medio di un processo stazionario in senso lato $X(t)$ è che valga

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T}^{T}|C_X(\tau)|d\tau = 0$$

8.17 È assegnato un segnale determinato periodico $p(t)$ di periodo T_0. Dire se il processo aleatorio $X(t) = p(t-\Theta)$, con $\Theta \in \mathcal{U}(-\pi,\pi)$, è ergodico in senso lato.

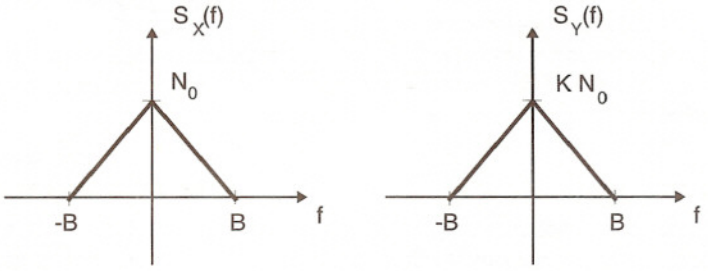


Figura 8.49

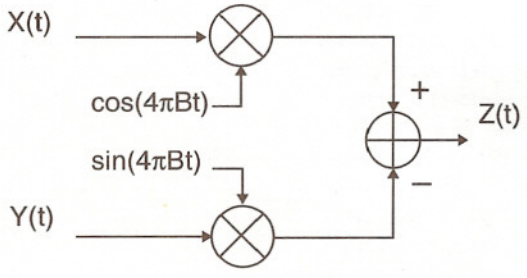


Figura 8.50

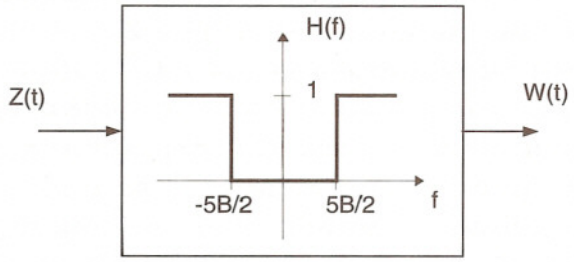


Figura 8.51

8.18 Dato un processo $X(t)$ ergodico in senso stretto con statistiche di qualunque ordine note, calcolare

$$\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\mathrm{u}[a-X(t)]dt$$

con a costante assegnata.

CENTRO BIBLIOTECARIO INGEGNERIA PISA